ISCP

INTERNATIONAL JOURNAL OF

COMPARATIVE PSYCHOLOGY

Sponsored by the International
Society for Comparative Psychology
and the University of Calabria

VOLUME 2, NUMBER 3

INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

Articles

INTRODUCTION: THE EVOLUTION OF COMPARATIVE PSYCHOLOGY

Nancy K. Innis

University of Western Ontario

Why do psychologists study the behavior of animals? The usual response to this question, even from many comparative psychologists, is that explaining animal behavior will help us understand human behavior, and understanding human behavior is the ultimate goal of psychology. Has this, in fact, been the aim of animal psychology, and if so, has it succeeded?

As scientific disciplines go, comparative psychology is relatively recent. Following the publication of Darwin's *Origin of Species* and his subsequent discourse on the evolution of intelligence (Darwin, 1859; 1871; 1872), psychologists became interested in studying animals and their relationship to humans. Early on, a fairly wide range of topics was explored and a large number of animal species studied. However, this was soon to change. The reader is referred to two excellent recent publications—Boakes' (1984) *From Darwin to Behaviourism* and Richards' (1987) *Darwin and the Emergence of Theories of Mind and Behavior*—which examine this early period.

The focus of inquiry narrowed, particularly in North America, when Behaviorism, in all its various forms, began to dominate psychology early in the twentieth century. Soon animal learning became the major area of inquiry. Evolutionary theory had paved the way for this development—if the differences between animal behavior and human behavior are simply differences of degree, then animals, whose behavior is supposedly simpler and more readily subjected to experimental control, are perfect subjects for the psychologist. Animal models of human behavior proliferated and one eminent theorist even "confessed" his belief that "everything important in psychology . . . can be investigated in essence through the continued experimental and theoretical analysis of the determiners of rat behavior at a choice point in a maze" (Tolman, 1938, p. 34).

While evolutionary theory permitted the development of animal psychology, studies of animal, and particularly human, behavior within an evolutionary framework were largely overshadowed by the seemingly successful enterprise of the animal learning theorists. With the study of representative species the future of comparative psychology began to

look bleak, and for decades now psychologists—and others—have been arguing about its impending demise. Recently, however, general process learning theory has fallen upon its own hard times.

At the meetings of the 24th International Congress of Psychology in Sydney, Australia at the end of August 1988, the International Society for Comparative Psychology was involved in the organization of a number of symposia that examined the mandate and current status of comparative psychology. In one of these, *Comparative Psychology: Towards the Year 2000,* convened by Ethel Tobach, speakers from countries around the world discussed the development and future of the discipline from their own national perspectives. Charles Tolman brought the session to a close on a positive note for comparative psychology, arguing that the methodology and evolutionary approach of this field may, indeed, provide a solution to the broader crises now faced by psychology in general. His paper follows on p. 197. Brian Mackenzie, whose paper appears on p. 189, also emphasized the importance of the evolutionary perspective of comparative psychology and its auspicious future in the symposium organized by John Barlow on *The Impact of Contingency Theory on the Development of Comparative Psychology.*

The symposium I organized for the Congress meetings was devoted to the question: *Has Animal Behavior Got Anything To Do With Human Behavior?* This session opened with a paper by Peter Livesey (Livesey & Bell, 1988) who described the ongoing program of research in his laboratory using a comparative method most closely identified with the work of Bitterman (1960, 1975; see also Mackintosh, 1974). Bob Boakes (1988) described work on the detection of causal relations in humans, which grew out of his previous studies with animals. John Staddon (1988) presented a provocative paper suggesting that studying animals merely as models of human behavior is unproductive, and that intelligence— whether it be human, animal or machine—is the proper subject matter for psychology. The paper by Shawn Lockery and Stephen Stich, which appears on p. 157, can be seen as illustrating the theme developed by Staddon in his conference presentation. John Tooby and Leda Cosmides examined two evolutionary perspectives—adaptation and phylogeny— suggesting that depending on the issues addressed one or the other may be the more fruitful in directing attempts to understand both human and animal behavior. Their paper is on p. 175. In the introductory paper which follows, prepared for this issue of the Journal, John Staddon and I address the question—What should comparative psychologists compare?—and proffer a response: compare mechanisms not behaviors or functions, and in so doing be concerned with the importance of theory.

REFERENCES

Bitterman, M. E. (1960). Toward a comparative psychology of learning. *American Psychologist, 15,* 704-712.

Bitterman, M. E. (1975). The comparative analysis of learning: are the laws of learning the same in all animals? *Science, 188,* 699-709.

Boakes, R. A. (1984). *From Darwin to behaviourism: Psychology and the minds of animals.* Cambridge: Cambridge University Press.

Boakes, R. A. (1988, August). Animal learning theory and human conditioning. In N. K. Innis (Chair), *Has animal behavior got anything to do with human behavior?* Symposium conducted at the meeting of the International Congress of Psychology, Sydney, Australia.

Darwin, C. (1859). *On the origin of species by means of natural selection.* London: John Murray.

Darwin, C. (1871). *The descent of man, and selection in relation to sex.* London: John Murray.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Livesey, P. J. & Bell, J. A. (1988, August). The animal-human continuum in learning. In N. K. Innis (Chair), *Has animal behavior got anything to do with human behavior?* Symposium conducted at the meeting of the International Congress of Psychology, Sydney, Australia.

Mackintosh, N. J. (1974). *The psychology of animal learning.* New York: Academic Press.

Richards, R. J. (1987). *Darwin and the emergence of evolutionary theories of mind and behavior.* Chicago: University of Chicago Press.

Staddon, J. E. R. (1988). Animal psychology: The tyranny of anthropocentrism. In P. P. G. Bateson & P. H. Klopfer (Eds.) *Perspectives in ethology. Vol. 8, Whither ethology?* London: Plenum.

Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review, 45,* 1-41.

WHAT SHOULD COMPARATIVE PSYCHOLOGY COMPARE?

N. K. Innis
J. E. R. Staddon

University of Western Ontario and Duke University

ABSTRACT: Scientific psychology is a search for the mechanisms that underlie behavior. Following a brief history of the comparative psychology of learning, we suggest that comparative psychologists should focus on mechanisms rather than performances, and provide an example of a simple, formal mechanism to illustrate this point.

The modern history of comparative psychology begins with Charles Darwin (1809-1882). Darwin's theory of evolution through variation and natural selection has provided a conceptual framework for psychology as well as for biology. As Tooby and Cosmides (p. 175) point out, the two parts to Darwinian evolution have led to two kinds of approach to the study of animal behavior: an emphasis on variation underlies the phylogenetic approach; an emphasis on natural selection underlies the study of adaptation. In this introductory paper we consider the comparative psychology of *learning*, perhaps the most complex behavioral adaptation and the topic most popular with psychologists. After a brief overview of some historical highlights, we focus on the issue of comparison. What have comparative students of animal learning chosen to compare? Has their work been successful? What should be compared?

HISTORICAL BACKGROUND

Several recent volumes (e.g., Boakes, 1984; Dewsbury, 1984; Richards, 1987) examine in detail the history of comparative psychology and Darwin's influence on it; we mention only a few of the more important early contributions. Following the publication of *The Descent of Man*

Address correspondence to either author at the Department of Psychology, Duke University, Durham, NC 27706.

(1871) and *The Expression of the Emotions in Man and Animals* (1872), in which Darwin's ideas concerning the evolution of intelligence were presented in some detail, naturalists became interested in studying the animal mind. One of the first to write extensively on the topic was George Romanes (1848-1894), a friend and champion of Darwin. As Romanes (1882) indicated in the Preface to *Animal Intelligence,* his aim was to offer a scientific textbook of comparative psychology that provided both facts about the levels of intelligence attained by various species, and a consideration of animal intelligence in terms of Darwinian theory. While Romanes' intentions were scientific, his reliance on anecdotes soon drew severe criticism, particularly from C. Lloyd Morgan (1852-1936). It was not long before Morgan (1894), in his own textbook *An Introduction to Comparative Psychology,* established one of the standards for future work in comparative psychology. In what came to be known as Morgan's Canon, he warned against anthropomorphic explanations, recommending that care should be taken not to interpret an action as the result of a higher mental process if, in fact, it could be explained by one at a "lower level in the psychological scale" (p. 53). Nevertheless, despite Lloyd Morgan's strictures, most comparative psychologists believed they were comparing faculties or abilities—not just phenomena or activities, and certainly not processes, in the modern sense.

The textbooks of comparative psychology at this time were in the tradition of contemporary works in comparative biology and anatomy. Chapters considered, in turn, examples of the behavior of a large number of species of various degrees of complexity and relatedness, on a variety of tasks. The emphasis then was on identifying and examining similarities and differences in capability and intelligence across animal species. The most influential textbook of this type in North America during the early years of the twentieth century was Margaret Washburn's *The Animal Mind* (1908), which was to remain popular through several subsequent editions. But this tradition was not to persist: While there would always be some biologically oriented students of animal behavior who maintained a concern with phyletic differences, the stream of psychology was soon to be diverted by a different current.

American Functionalism—Mind Surrenders to Behavior

A new theoretical force—Behaviorism—emerged early in the twentieth century. Although eventually it was to radiate into many varieties, from radical to purposive, Behaviorism from the outset altered the approach taken by comparative psychologists. Behaviorism was the offspring of the functional movement that was gaining ascendance in North America at this time, particularly at places like Chicago and Columbia. Since the predominant aim of the functionalists was to ameliorate the human condition, it is not surprising that they found the

adaptive function of behavior more intriguing than its evolutionary origins, and situational factors more significant than hereditary predispositions. As this new wave of environmentalist thinking took hold, boosted by the polemical efforts of John B. Watson (1878–1958), the "father"—or at least the publicist—of Behaviorism, evolutionary theory provided the justification for studying animal behavior as the precursor of the behavior of human beings. But evolution was otherwise largely ignored.

As the name Behaviorism implies, animal mind was displaced by animal action. Most experimental psychologists began to study the ways in which animal behavior was altered by environmental experience— animal learning. Two basic approaches to learning, initially ill-distinguished, soon began to shape the study of animal behavior, as Pavlovian conditioning and Thorndike's Law of Effect became the central paradigms guiding animal research. With these developments, researchers soon shifted away from studying a wide range of species, looking instead at the details of learned responding in a small, but presumably representative, set. The fields of comparative and animal psychology to a large extent were subsumed by learning theory. This change is exemplified by the emphasis placed on animal learning in one of the most popular texts, Moss's (1934) *Comparative Psychology,* a book that was widely used through three editions over the next 25 years. There were, of course, exceptions to these general trends, particularly in the work of Schneirla and his associates (see, for example, Maier & Schneirla, 1935/1964), and there is some room for difference of opinion about the dominance or otherwise of learning theory during this period (cf. Dewsbury, 1984; Innis, 1987).

Nevertheless, it seems fair to summarize the recent history of comparative psychology in terms of the shift from a concern with phylogeny to a predominant interest in learning and the effects of experience. Following Darwin, students of animal intelligence attempted to identify the intellectual capacities of various animal species and to categorize them in ways that might say something about the evolutionary history of the species and the capacity studied. This phyletic approach was purely descriptive and initially involved primarily anecdotal evidence, although later more scientific field studies were carried out. A major change in approach occurred, particularly in North America, as the functionalist, and later behaviorist, schools of psychology began to predominate. Animals now became tools, in contrived laboratory settings, with which to examine a small set of general rules proposed by learning theorists. These rules concerned the ways in which behavior changes as the result of experience and were of interest primarily for what they could teach us of ways to improve human life. The focus shifted away from phyletic comparison towards the prediction and control of action, with clues to better technique to be provided by limited cross-species comparison.

The Comparative Psychology of Learning

It was not just that psychologists turned to the study of animals to explain the behavior of humans. They also accepted, often without much discussion—at least during the early years—the idea that associative learning is a primitive property, a basic element of the vertebrate *bauplan*, like quadrupedalism or the circulatory system. Regarded from this point of view, the differences between ape and reptile—and man—could be seen as mere parameter adjustments and complexifications of the same basic general learning process. This general process assumption allowed animal psychologists to focus on a small set of representative species—monkeys, pigeons, but primarily the white rat—for which experimental procedures had been well-tested and whose maintenance conditions were well-known.

In an attempt to provide a fresh agenda for comparative psychology M. E. Bitterman (1960; 1975), a Schneirla student in his undergraduate days, suggested a new, more sophisticated goal: Rather than looking just at the ability of different species to perform adequately on supposedly comparable tasks—an almost impossible constraint, given vast and hard-to-quantify differences in motivation and sensory, motor and perceptual abilities among species—comparative psychologists should be comparing the functional relationships displayed by different species under roughly analogous learning conditions. For example, while we cannot reasonably expect a goldfish to press a target as fast or as forcefully as a rat, perhaps the relation between (say) rate of pressing and reward rate will be similar in both species.

Bitterman concentrated on data for five "prototype" species (goldfish, turtle, pigeon, rat and monkey), examining a number of learning phenomena that had been extensively studied in rats: reversal learning, probability matching, partial reinforcement effect, reinforcement contrast effects, and several others. He reported numerous similarities across species, and also a few notable differences. For example, he claimed that fish and rats differ in the processes underlying both serial reversal and probability learning. Pigeons he found to be rat-like (they improved) on serial reversal, but fishlike (they matched rather than maximized) on some probability learning tasks (Bitterman, 1965). His conclusion that these empirical differences between a "lower" and a "higher" species must reflect the progressive evolution of new mechanisms did not go unchallenged, however. Mackintosh (1969; Bitterman & Mackintosh, 1969), while not questioning the differences in performance, took issue with the inference that different underlying processes are involved. Mackintosh maintained that the behavioral differences may be merely quantitative and could be accounted for by a single mechanism, "say, some differences between the parameters used to explain the behavior of rats and those used for the fish . . ." (1969, p.

138), a position identified particularly with the earlier work of C. L. Hull (e.g., 1943). Unfortunately, Mackintosh made no concrete suggestions along these lines, referring instead to an attentional hypothesis that he himself characterized as "vague," which was not convincing to Bitterman. Thus, the one- vs. two- process dispute remained unresolved.

The apotheosis of the one- process view is the recent work of Macphail (1982; 1987) who has argued (somewhat extravagantly) for the essential identity of all nonlinguistic learning across the whole vertebrate phylum—although his arguments still lack the kind of quantitative theoretical basis identified as ideal by Hull and Mackintosh.

Research along the lines proposed by Bitterman has focused recent attention on the question of learning processes. Nevertheless, the fact that Macphail (1987) can seriously argue for the essential identity of all nonlinguistic intelligence, and elicit a mass of respectful if largely critical commentary, shows we still lack deep understanding of the processes involved in associative learning. Despite the efforts of Bitterman, Mackintosh and a few others, the experimental comparative focus is still very much on phenomena rather than mechanisms. Perhaps it is time to put some flesh on the suggestion that many, perhaps all, differences among species can be accounted for by "mere" parameter variation of a single, or at most a small number, of "basic processes."

MECHANISMS OF ANIMAL LEARNING

Our thesis, which we will illustrate by example, is that assertions about the essential identity, or lack thereof, of any faculty or process across species are vacuous in the absence of fairly detailed specification of the mechanisms involved. The papers by Bitterman and Mackintosh in which this issue first came to the fore are now quite old, but process equivalence is far from a dead issue in psychology. It is a hot topic in memory research for example, where vigorous debate surrounds the question of multiple memory systems: How many are there, and what are their properties (e.g., Tulving, 1985; Sherry & Schacter, 1987)? In comparative psychology, Macphail's (1982; 1987) provocative thesis has fueled recent controversy about the essential unity or diversity of all infrahuman intelligence. Bitterman, for example, has taken strong issue with Macphail's evaluation of the *successive negative contrast effect* (SNCE), i.e., the slower or less accurate responding of animals trained with a small reward following training with a larger reward, compared with the performance of animals trained from the outset with a small reward. Goldfish generally fail to show this effect, which is easily demonstrated in rats and several other animals and birds. On this basis, Bitterman concludes that goldfish are lacking some mechanism that evolved later on: ". . . the mechanism of successive negative contrast (whatever it

may be) evolved in some common reptilian ancestor of birds and mammals (Bitterman, 1987, p. 659)." Macphail, on the contrary, argues that the failure to find the SNCE in goldfish is not decisive disproof for his unitarian position.

We believe that this argument—indeed, all such arguments—cannot be resolved without putting in play some specific theory, either for two processes (Bitterman's position) or one (Macphail's position). We will try to illustrate our point by example. We hope we can persuade comparative psychologists to relinquish fruitless and inconclusive debates of the "one mechanism or two" variety, and concentrate instead on proposing and testing specific, detailed mechanisms for the phenomena in which they are interested.

Habituation and Sensitization: A Simple Model System

We are not yet in a position to offer the kind of theory of the SNCE that would satisfy Mackintosh's criterion. But it is possible to speculate in a relatively straightforward way on mechanisms for the much simpler phenomena of habituation and sensitization. Our aim is not so much to provide a definitive theory as to make a methodological point using these elementary adaptive phenomena as an example.

Habituation is the waning in strength of response, such as the startle response, to repeated stimulation. Sensitization is an increase in the strength of a response as a stimulus is presented repeatedly. Figure 1

Time
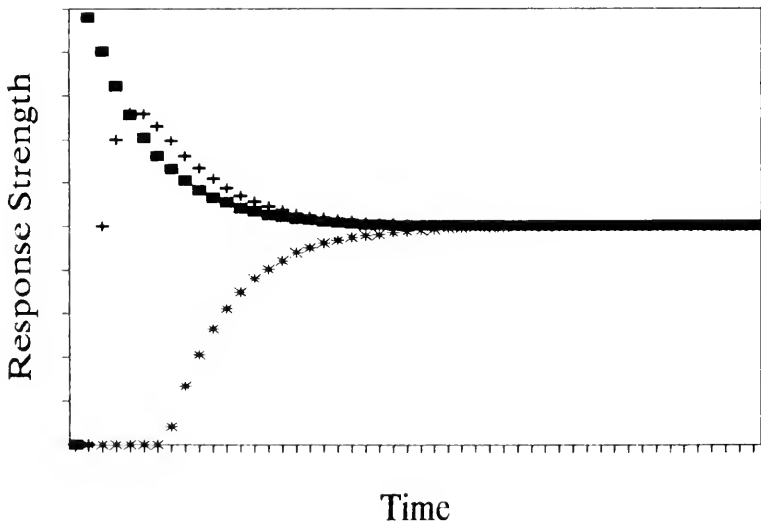
FIGURE 1. Response of three hypothetical "species" to repeated stimulation. Species A (filled squares) shows habituation; species B (stars) shows temporal summation and sensitization; species C (crosses) shows a biphasic response. Parameter a is a time constant: values for each species are given. See text.

shows three sets of imaginary data from habituation/sensitization experiments with stimuli of comparable affective tone. The three graphs show response strength, of three species, to a series of repeated stimuli. Species A (filled squares) shows clear evidence for habituation: response strength decreases across the series. Species B (stars) shows temporal summation and then sensitization: initial stimulus presentations have no effect, but then response strength increases across the series. Species C (crosses) shows a biphasic effect: at first there is sensitization—response strength increases—but then there is habituation (response strength declines from its maximum value early in the series).

The question: Can we tell from these data whether there is a single underlying process across all three species, or two processes, or several? Obviously we would want to do more experiments before committing ourselves. Figure 2 shows the results of one such experiment, in which we have varied intertrial interval, which was one time unit in Figure 1, and is two units in Figure 2. Now we see larger differences among the three species. Species A still shows habituation, species B sensitization and species C a biphasic effect, but the absolute levels of the responses now differ substantially among the three species: species A consistently shows the highest level of responding, B the lowest and C is intermediate. Moreover, species A's steady-state response has increased, but the steady-state response of the other two species has decreased, an informative dissociation, one might think. Evidently trial spacing affects the absolute level but not the general form of the response vs. trials function.
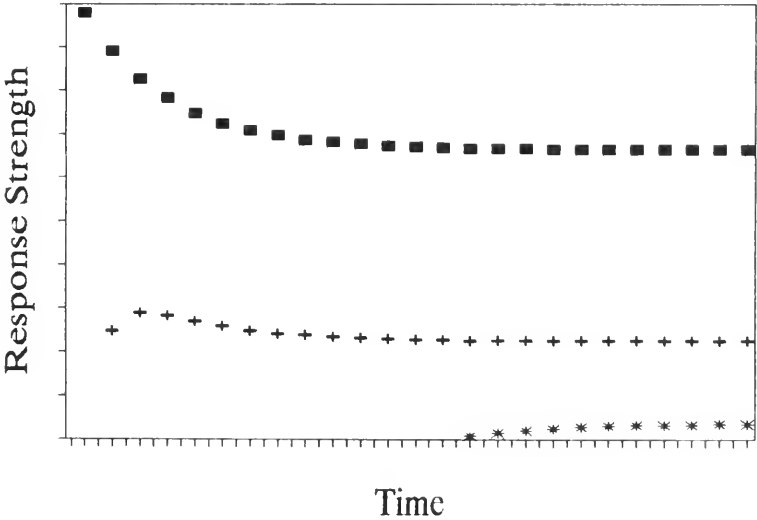
FIGURE 2. Response of three hypothetical species to repeated stimulation at an inter-stimulus interval twice that in Figure 1. Species A; filled squares; species B: stars; species C; crosses. Parameter a, as in Figure 1.

Given this kind of stability of pattern differences, but divergences in direction of change under experimental manipulation, many would feel safe in concluding that here we do indeed have at least two and perhaps three different processes: Species A (a fish, perhaps) shows process 1 (habituation), species B (reptile?) process 2 (sensitization), and species 3, which is clearly more advanced (mammal?), both processes.

Figure 3 shows the process that actually generated all the curves in the first two figures. It is a single process, by definition, but it does have several identifiable components: two integrators, a comparator and a threshold. Each integrator is a "leaky bucket" process that can be described by a simple discrete-time linear operator. Thus, for integrator 1 (which is responsible for both temporal summation and sensitization):

$$Ve(t + 1) = aVe(t) + (1\text{-}a) X (t) \qquad (1)$$

where $Ve(t+1)$ is the output of the system at time $t+1$, $X(t)$ the stimulus input at time t, and a is a time constant: the higher the a value, the smaller the effect of the current input compared to past inputs, i.e., the longer the time period over which temporal summation can occur. A similar equation describes the second integrator (which accounts for habituation).

$$Vi(t + 1) = bVi(t) + (1\text{-}b)Vo(t) \qquad (2)$$
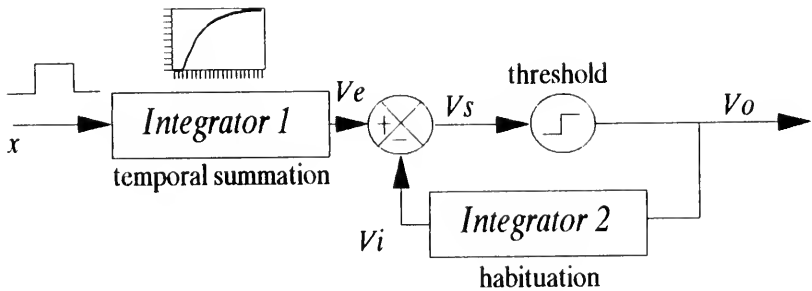
FIGURE 3. System of two integrators and a threshold that produced the data shown in Figures 1 and 2. Parameter values: a = 0.01 (species A), 0.9 (species B), 0.5 (species C); b = .09 and θ = 0.5 for all three species. See equations in text.

Two additional equations complete the picture

$$Vs(t) = Ve(t) \text{ - } Vi(t) \qquad (3)$$

$$Vo(t) = V_S(t) \text{ - } \theta, \text{ if } V_S > \theta, 0 \text{ otherwise} \qquad (4)$$

Thus, the system is completely described by four equations with three free parameters: two times constants, a and b, and a threshold, θ. Equations 1 and 2 are integrators; Eq. 3 describes how Ve (the excitatory effect of the stimulus) and Vi (the inhibitory effect of the response—a good Hullian concept) combine to produce the "action potential," Vs. Equation 4 describes how action potential, Vs, is "thresholded" to produce response amplitude, Vo.

The relation among the three main variables in the system, Ve, Vi, and Vs, and the input, X, is shown graphically in Fig. 4.
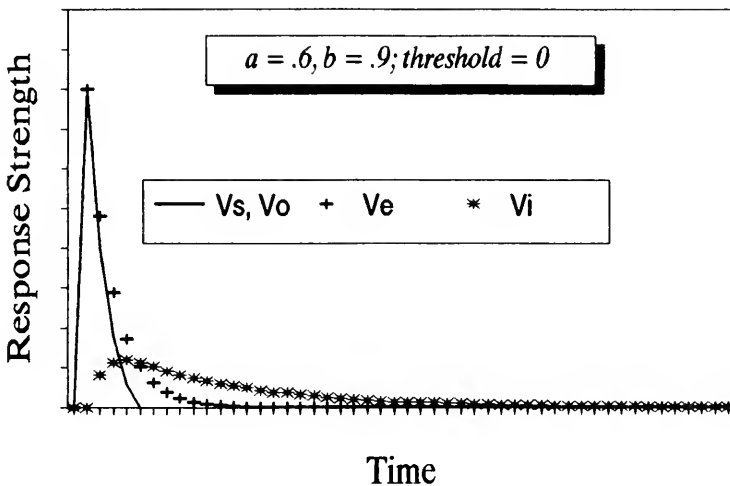
FIGURE 4. Response of the components of the system shown in Figure 3 and equations 1-4 in the text to a single stimulus presented at t = 0. The threshold parameter, θ, is set to zero; time constant $a = 0.6$, time constant $b = 0.9$. Solid line: system output ($V_s = V_e V_i$' = V_o' for $\theta = 0$); crosses: exitatory response (first integrator: V_e); stars: inhibitory response (second integrator: V_i).

Thus, despite the apparent differences among our three species, a relatively simple process, determined by just three parameters, is sufficient to generate all of them. Moreover, the differences among species reflect variation in just one of the three parameters, a, the time-constant for the excitatory (sensitization) component. We can also say that the system described in Eqs. 1-4 is close to the minimal system, in the sense that an asymmetrical biphasic response pattern, the most complex pattern in the data, requires at least two parameters for its specification. But of course there are very many other two- or three-parameter systems that might be adequate to account for this limited data set.

To validate this or any other black-box model, two further steps are necessary. First, test experimental predictions: a real-time dynamic model like this is very powerful, in the sense that it makes exact predic-

tions about the pattern of response amplitude to any pattern of input stimuli. It is therefore exceedingly easy to disconfirm (perhaps this is why so few such models survive). Second, if the model withstands all experimental tests, it becomes worthwhile to look beyond the *formal* mechanism to the actual *physiological* processes that must carry out the formal operations (cf. Fentress, 1980). Can we find measurable physiological processes (transmitter uptake rates, creation or dissolution of ionic channels, neuroanatomical constraints, etc.) corresponding to the structure and parameters of the model? This second step is more likely to be successful the better the model judged purely on the basis of black-box experiments. Moreover, step 1 need not be absolutely complete before embarking on step 2. Once the main outlines of the formal process are clear, it may well be more efficient to refine the details by looking closely at the neurophysiological underpinnings—than by further black-box elaborations. No one can provide a recipe for the optimal research strategy; each problem must obviously be judged on its own merits and in light of the information and research opportunities available at each level, blackbox or physiological.

What is the role of between-species comparison in this endeavor? We believe that variation in species is not different—in its role as part of a research strategy—from variation in any other experimental condition, such as motivational type and level, stimulus intensity and quality, stimulus sequence, etc. We know something, *a priori*, about each of these things. Stronger stimuli we know will generally elicit larger responses from the appropriate sensory receptor and will be subject in their effects to psychological constraints such as Weber-law effects. Similarly, different species stand in more or less well-known phylogenetic relationship to each other. Species A may be closer in its morphology to the common ancestor of species A, B and C than either B or C; thus we may expect that many of the features of B and C represent identifiable transformations (according to the kinds of developmental rules discussed by Gould, 1977, for example) of homologous features in C. Although there are no set rules for comparison, these phylogenetic relationships impose constraints that may limit the search for possible mechanisms to sets that by simple transformation or parameter adjustment can encompass the behavior of all three species. Thus, following the only true scientific method—of guesswork, donkeywork, prediction and test—between species comparisons can aid the search for understanding of behavior, brain, and the relations between the two.

CONCLUSION

Scientific psychology is a search for the mechanisms that underlie behavior. Every scholarly field, as part of its dynamic of self-preservation,

tends to be both imperialistic and exclusionary (Staddon, 1987). Thus, we have economists seeking to explain biology or psychology in economic terms, literary "deconstructionists" arguing that scientific theories are neither true nor false but merely "texts" as subject to their interpretation as *Fanny Hill* or *The Scarlet Letter* (see Crews, 1986, for a critical account)—and psychologists asserting that there are laws of individual behavior that transcend, or at least cannot be reduced to, physiology. While this view may ultimately be borne out, we see little reason to accept it on present evidence. The best worked out psychological laws, of color vision or psychophysics, for example, have always turned out to relate very directly to the neural machinery involved.

From this point of view, between-species comparison is just one of a number of tools we can use to understand mechanism—formal or, ultimately, physiological—and as we have tried to show, many comparative psychologists, such as Mackintosh and Bitterman, have so viewed it. Like any tool, its value is to be judged by its effectiveness: How helpful is it to our understanding of learning mechanisms (for example) to compare the learning of goldfish, rat and pigeon? Since we still know rather little about the learning mechanisms of any species, this question is still open—although few would doubt the value of looking at more than one species.

We have argued that there are important limitations on what can be inferred from comparing learning performance across species: Different performance patterns cannot be taken as evidence for different underlying mechanisms. And, symmetrically, similar performance cannot be taken as evidence for similar mechanisms (the monotonic 2-parameter curves in Fig. 1 and 2 could easily be modeled by a process with one parameter plus a constant base rate, for example). These limitations are in fact quite well known to most thoughtful theorists (cf. Elepfandt, 1987). Nevertheless, the obvious implication is widely ignored: To argue for one, two, or many learning processes in the abstract, without specifying what these processes actually are—at a level where they can be subjected to the kind of searching test we illustrated with our habituation model—is to embark on a debate without any possibility of resolution.

ACKNOWLEDGEMENTS

REFERENCES

Bitterman, M. E. (1960). Toward a comparative psychology of learning. *American Psychologist, 15,* 704-712.

Bitterman, M. E. (1965). Phyletic differences in learning. *American Psychologist, 20,* 396-412.

Bitterman, M. E. (1975). The comparative analysis of learning: are the laws of learning the same in all animals? *Science, 188,* 699-709.

Bitterman, M. E. (1987). Evidence of divergence in vertebrate Learning. *Behavioral and Brain Sciences, 10,* 659-660.

Bitterman, M. E. & Mackintosh, N. J. (1969). Habit-reversal and probability learning: Rats, birds and fish. In R. M. Gilbert & N. S. Sutherland (Eds.) *Animal discrimination learning.* New York: Academic Press.

Boakes, R. A. (1984). *From Darwin to Behaviourism: Psychology and the minds of animals.* Cambridge: Cambridge University Press.

Crews, F. (1986). *Skeptical engagements.* New York: Oxford University Press.

Darwin, C. (1871). *The descent of man and selection in relation to sex.* London: J. Murray.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: J. Murray.

Dewsbury, D. A. (1984). *Comparative psychology in the twentieth century.* Stroudsburg, PA: Hutchinson Ross.

Elepfandt, A. (1987). Comparative cognition: Inadequate approach, precipitate conclusions. *Behavioral and Brain Sciences, 10,* 661-662.

Fentress, J. C. (1980). How can behavior be studied from a neuroethological perspective? In H. M. Pinsker & D. W. Willis, Jr. (Eds.) *Information processing in the nervous system.* New York: Raven Press.

Gould, S. J. (1977). *Ontogeny and phylogeny.* Cambridge, MA: Belkap Press.

Hull, C. L. (1943) *Principles of behavior: An introduction to behavior theory.* New York: Appleton-Century.

Innis, N. K. (1987). Edward C. Tolman: Comparative psychologist? In E. Tobach (Ed.) *Historical perspectives and the international status of comparative psychology.* Hillsdale, NJ: Erlbaum.

Mackintosh, N. J. (1969). Comparative studies of reversal and probability learning. In R. M. Gilbert & N. S. Sutherland (Eds.) *Animal discrimination learning.* New York: Academic Press.

Macphail, E. M. (1982). *Brain and intelligence in vertebrates.* Oxford: Clarendon Press.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behavioral and Brain Sciences, 10,* 645-695.

Maier, N. R. F. & Schneirla, T. C. (1935/1964) *Principles of animal psychology,* Enlarged Edition. New York: Dover.

Morgan, C. L. (1894). *An introduction to comparative psychology,* London: Walter Scott.

Moss, F. A. (Ed.). (1934). *Comparative psychology.* New York: Prentice-Hall.

Richards, R. J. (1987). *Darwin and the emergence of evolutionary theories of mind and behavior.* Chicago: University of Chicago Press.

Romanes, G. (1882). *Animal intelligence.* New York: Appleton.

Sherry, D. F. & D. L. Schacter (1987). The evolution of multiple memory systems. *Psychological Review, 94,* 439-454.

Staddon, J. E. R. (1987). Science and pseudoscience. *Interdisciplinary Science Reviews, 12,* 114-116.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385-398.

Washburn, M. F. (1908). *The animal mind: A text-book of comparative psychology.* New York: Macmillan.

PROSPECTS FOR ANIMAL MODELS OF MENTAL REPRESENTATION

Shawn Lockery and Stephen Stich
University of California, San Diego

ABSTRACT: A major goal of physiological psychology is to determine the physical basis of mental representation. Animal models are essential to this project. Dretske's influential analysis of the concept of mental representation suggests that operant and classical conditioning involve mental representation. This analysis comports well with known physiological mechanisms of conditioning, but fails to capture necessary features of mental representation at the human level. We conclude that the applicability of animal models to the problem to human mental representation is more restricted than previously thought.

What is the relationship between the study of animal psychology and the study of human psychology? Staddon (1988) reminds us that there is a long tradition that answers this question by appealing to the idea that in various psychological domains, animals can be used as models for people. And, of course, in this tradition it is human psychology that is ultimately of interest. Staddon is not comfortable with this tradition. He argues that it has been a baleful influence on the study of animal psychology while yielding relatively little insight into human psychology. The alternative that Staddon recommends is that "psychology as a basic science should be about intelligent and adaptive behavior, wherever it is to be found," and thus that animals should be "studied in their own right for what they can teach us about the nature and evolution of intelligence, and not as surrogate people, or tools for the solution of human problems." This is a view we wholeheartedly endorse. Much of what we say in this paper can be read as illustrating Staddon's theme.

The domain in which we propose to develop our illustration is the rich literature on intentionality and mental representation. Our thesis will be that much recent theorizing about the concept of mental representation and its role in psychological explanation has been led astray by its anthropocentric focus. Most theorists simply presuppose that mental representation is a single phenomenon, and that the paradigm case of the phenomenon is to be found in conscious human thought. Typically, the psychological processes of animals are seen as relevant to the understanding of intentionality only in so far as they provide simple models for

---

full-blown human intentionality. It is our contention that this anthropocentric approach is in two ways unfortunate: first, it yields relatively little insight into human intentionality, and second, it obscures the fact that there may be very different kinds of semantic or intentional phenomena and thus that significantly different notions of representation may all have a substantive role to play in the explanation of intelligent and adaptive behavior. The specific target of our critique will be the elegant and sophisticated account of mental representation developed by Fred Dretske (1988) in his recent book, *Explaining Behavior.*

This paper is divided into three parts, the first of which is largely expository. In it we sketch Dretske's account of mental representation, and the role that it plays in psychological explanation. Dretske's story is developed against the background of some abstract and very schematic assumptions about what goes on in certain sorts of conditioned learning. So a natural question to ask is how well Dretske's cartoon of conditioning comports with what is known about the underlying neurobiology. This is the project pursued in the second section. The answer is that despite its simplifications, Dretske's sketch of learning meshes quite well with the emerging neurobiological details. In the third section, our question will be how useful Dretske's account of mental representation is likely to be for the understanding of the paradigm cases of human intentionality. Our contention is that Dretske's account is a poor model for the sort of intentionality and intentional explanation that loom large in human psychology. On the brighter side, however, we argue that Dretske has isolated a semantic or representational notion that can be of use in animal psychology. If this is right, then it is to be welcomed on its own merits, not disparaged because it fails as a model of human intentionality.

AN OVERVIEW OF DRETSKE'S PROJECT

The subtitle of Dretske's book is *Reasons in a World of Causes,* and that subtitle provides a convenient jumping off point for our description of Dretske's project. Following tradition, Dretske begins with a discussion of human behavior and its explanation. Fred, who is sitting in the living room, gets up and walks into the kitchen. How can this behavior be explained? Common sense psychology often provides a ready answer. Fred walked into the kitchen because he wanted a drink, and he believed he could get one there. This explanation, which provides Fred's reasons for walking into the kitchen, invokes a pair of intentional or representational states. His belief represents the world as being in a certain way. (Had Fred believed, instead, that there was nothing to drink in the kitchen, he would have remained in the living room, or gone to search

elsewhere). And his desire has as its object some future state—Fred getting a drink—which may, or may not, ultimately come to pass.

But it is also the case that when Fred walked into the kitchen, he did so because various muscles contracted; they, in turn, were "responding to a volley of electrical impulses emanating from the central nervous system." (p. ix). Walking into the kitchen was something Fred's body did, and ultimately we expect that the movements of his body will be explained by neuroscience and biology. If this is right, Dretske notes, then "one seems driven, inevitably, to the conclusion that, in the final analysis, it will be biology rather than psychology that explains why we do the things we do" (1988, p. x). "What, then," Dretske asks,

> . . . remains of my conviction that I already know, and do not have to wait for scientists to tell me, why I went to the kitchen? I went there to get a drink, because I was thirsty, and because I thought there was still a beer left in the fridge. However good biologists might be, or become, in telling me what makes my limbs move the way they do, I remain the expert on what makes me move the way I do. Or so it must surely seem to most of us. To give up this authority, an authority about why we do the things we do, is to relinquish a conception of ourselves as human agents. This is something that we human agents will not soon give up (p. x).

The ultimate goal of Dretske's project is to preserve a legitimate role for intentional psychology by showing how two schemes for explaining behavior—the intentional (or "psychological," as Dretske sometimes says) and the neurobiological—can co-exist. He wants to find an explanatory role for reasons in a world of causes.

A central element in Dretske's account of the explanatory role of reasons is his distinction between triggering causes and structuring causes. Before explaining that distinction, we need to think a bit about the relation between behavior and bodily movement. Most behavior involves bodily movement: Fred reaches for a beer; Bonnie turns the steering wheel of her car to avoid hitting a child; a rat depresses a lever with its paw when the light in its cage goes on in order to get a bit of food. However, not all bodily movement will plausibly count as behavior. When Clyde pushes Bonnie's arm, or I move the rat's paw, the resulting bodily movements are not part of Bonnie's behavior or the rat's. The distinction, Dretske urges, turns on the location of the cause of the movement. In genuine cases of behavior, a salient aspect of the cause of the movement is some event or process internal to the organism. By contrast, when I move the rat's paw the salient causes are external to the rat. In the terminology Dretske recommends, the term "behavior" is reserved for a process in which some internal state or event causes a bodily movement —schematically, a behavior is a process in which an internal state C causes a movement M. The movement itself is the visible product of the behavior; it is the output of the process.

Now when behavior is construed as a process, the question of why a particular piece of behavior occurred can be construed in two very different ways. On one reading it is a request for some explanation of what began the process and kept it going. This is what Dretske calls a triggering explanation. In the case of the rat pressing the lever, the explanation would begin with the light. It would detail how the light brings about some internal perceptual state C, and how C leads to the production of certain movements of the rat's paw. Ultimately, one would hope to be able to tell the entire story, from retinal stimulation to paw movement, at the neurobiological level. However, there is another aspect of the rat's behavior that needs to be explained. In addition to asking how a certain stimulus brings about an internal perceptual state, and how that internal state leads to certain movements, we can ask why the organism is so structured that the internal state leads to those movements rather than others. Why is C connected to M, rather than to M'? In asking this question, we are seeking what Dretske calls a structuring explanation, rather than a triggering explanation.

Since the distinction between triggering and structuring explanations is central to Dretske's account, we'd do well to consider another case in which the contrast emerges quite vividly. Suppose we are sitting at a friend's home and, as the day gets warmer, suddenly his automatic garage door motor is turned on, and the garage door opens. Why did this happen? Well, the rising temperature bent the bimetallic strip in the thermostat on the wall. That closed a circuit, which enabled current to flow to the garage door motor. (Note the pattern: an environmental E caused an internal C which caused the movement M.) The temperature triggered the opening of the door, and with a bit of effort we could tell an elaborately detailed story about the physical processes that subserve each of these causal connections. Having been told all of this, however, there is still something important that needs to be explained. Why is the thermostat connected to the garage door? Why isn't it connected to the furnace, or the air conditioner, or the doorbell, for that matter? Here, of course, all sorts of answers are possible. Perhaps our friend has found that opening the garage door cools off the house. Perhaps it is a way of letting the dog out in warm weather. Perhaps the thermostat was wired to the garage door as a joke. Whatever the explanation, it will be very different from the triggering explanation, because what we want explained is not the mechanism by which an environmental event leads to a certain movement. What we want to know is why the system is structured in this way rather than in some other way.

Recall that Dretske's fundamental problem is to find a place for intentional (or semantically interpretable) phenomena in the explanation of behavior. How could the fact that an internal state means some-

thing, or represents the world as being a certain way, contribute to the explanation of behavior? The answer that Dretske proposes focuses on structuring causes. Certain internal states of organisms lead to certain movements—the organisms are hooked up in that way—because they represent a particular fact or state of affairs. To see how this sort of explanation works, we need two further notions; indication and representation.

Indication, for Dretske, is the basic building block out of which more complex semantic or representational notions are built. One state of affairs indicates another if the occurrence of the latter is strongly correlated with the occurrence of the former. Thus, for example, in certain sorts of trees, the fact that the fifth ring in a cross section of the trunk is significantly wider than the other rings indicates that the tree grew more vigorously in its fifth year of life than in any other year. This notion of indication has a pair of important features. First, indication is a perfectly naturalistic notion; there is nothing spooky or mysterious about it. Second, indication is a very promiscuous relation. One state of affairs can be an indicator for many others. Thus, for example, if the tree whose rings we are examining lived in an arid environment, the larger than average fifth ring may also indicate that rainfall was significantly above average in the fifth year of the tree's life.

Let's turn now to representation. For Dretske, representation is a relation that obtains between an indicator and a state of affairs when the indicator has the function of indicating that state of affairs. Here again, nonbiological examples provide intuitive illustrations. In a thermostat, the degree of curvature of the bimetallic strip indicates the ambient temperature. Moreover, in the thermostat, it is the function of the strip to indicate the temperature. In this case it is easy enough to be very explicit about the function of the bimetallic strip, since an engineer designed the thermostat with just that indicator function in mind. It was because the strip indicates the temperature that the engineer wired it up in the way he did. Note here that we have a structuring explanation which appeals to the protosemantic indicating function of the bimetallic strip. An explanation of why the thermostat is hooked up in the way it is invokes the fact that the bimetallic strip indicates temperature.

Now, of course, there is a sense in which this is not a terribly interesting case of a structuring explanation, if our ultimate aim is to understand the explanatory role of representations, since the structuring is done by a person, whose beliefs, goals and other intentional states remain unexplained. But a central step in Dretske's attempt to find an explanatory role for intentional phenomena is his contention that there are phenomena to be found even in relatively simple organisms whose explanation is structurally analogous to the one just given for the ther-

mostat. There are biological processes in which an internal indicator comes to be hooked up to a movement controlling mechanism because of what it indicates.

The paradigm for Dretske's semantically involved structuring explanations is provided by operant conditioning. In operant conditioning, Dretske maintains, an internal indicator comes to have causal control over a movement producing mechanism: a C gets linked to an M. Moreover, this happens because the C is an indicator of some environmental feature. Consider the example of the rat in the Skinner box. At the beginning of the rat's training, a light goes on, and there is some internal state of the rat that indicates this fact. (Indeed, there are probably lots of different internal states that indicate the light being on. Indication, recall, is just lawful correlation). But at the outset none of the internal indicator states cause the rat to depress the lever. Sooner or later, however, the rat will happen to depress the lever when the light is on, and the result will be a reward. Note that the reward is contingent on two things: the light must be on and the lever must be pressed. Under these circumstances, some internal indicator of the light being on, some C, will gradually come to cause a kind of bodily motion it did not cause before. If it were not for the fact that the reward was causally contingent on the light being on, the internal indicator of the light would not have ended up linked to the movement-producing mechanism. So a semantic property of C—the fact that it is an indicator of the light being on—plays an essential role in a structuring explanation. If we want to know why C is hooked up to M—why the system is structured in this way—then the fact that C indicates the light being on is going to be an important element in the explanation. The pattern of this explanation is indicated schematically in Figure 1.
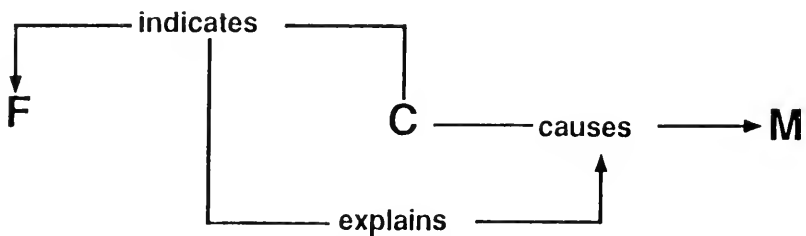
FIGURE I. Causal and explanatory relations underlying indication and representation (after Dretske, 1988, p. 84). F is an event or condition of the environment and C the internal state of the organism which "indicates" that F occurs. In addition to indicating F, C comes to "represent" F, when part of the explanation of the causal link between state C and movement M is the fact that C indicates F.

"Once C is recruited as a cause of M," Dretske maintains,"—and recruited as a cause of M because of what it indicates about F—C acquires, thereby, the function of indicating F. Hence, C comes to repre-

sent F. C acquires its semantics, a genuine meaning, at the very moment when . . . the fact that it indicates F . . . acquires an explanatory relevance" (p. 84). Moreover, according to Dretske, once C is hooked up to M in this way, it acquires the status of a genuine belief (or proto-belief). That is because, in Dretske's view (borrowed from Ramsey 1931 and Armstrong 1973), a belief is an internal map "by means of which we steer" (p. 79). A bit less metaphysically, "beliefs are representational structures that acquire their meaning, their maplike quality, by actually using the information it is their function to carry in steering the system of which they are a part" (p. 81). "What you believe, i.e., the semantic content of your belief is relevant to what you do because beliefs are precisely those internal structures that have acquired control over output, and hence become relevant to the explanation of system behavior, in virtue of what they, when performing satisfactorily, indicate about external conditions" (p. 84).

FINDING THE C's AND THE M's: DRETSKE MEETS NEUROSCIENCE

Dretske's formulation of learning offers a plausible way of making explicit how semantic relations enter our explanations of behavior. Clearly, however, his account makes some rather demanding assumptions about the underlying neurobiology. Should any of these turn out to be unwarranted, Dretske's attempt to show how intentional psychology and neurobiology can co-exist would be undermined. We might wonder, for example, whether the neurobiologists find anything like Dretske's C's (explicit, identifiable indicators of environment conditions) or his M's (brain sites responsible for particular behaviors). An equally important question is how C's and M's become connected as the animal is conditioned. Can we really maintain that a C–M connection is established *because of what C indicates?* How are we to phrase this in neurobiological terms?

Ideally, to answer these questions, one would like to have a complete understanding of how operant conditioning—Dretske's example—works at the neurobiological level in a wide range of species. One could then determine whether it operates in general the way his account of the explanatory role of semantics supposes. Unfortunately, the neurobiology of learning and memory is still in its infancy and many decades away from what we require as far as operant conditioning is concerned. Nevertheless, at least one general mechanistic principle has begun to emerge. And while we may know little about operant conditioning, the situation is considerably better for other forms of learning, particularly in invertebrate species where these questions can be addressed somewhat more directly than in the human or even mammalian cases.

A central question for Dretske of course is whether indicators exist in the nervous system and whether these function during learning as he would suppose. We have known for a long time that states of the central nervous system co-vary with environmental stimuli and so "indicate" them. This issue has been elegantly explored by Mountcastle and coworkers (1957) in the somatosensory system and in the well known studies of Hubel and Wiesel (1977) in the visual system. Thus, it is practically part of the neurobiological canon that something like Dretske's C's do exist. Roughly the same can be said for movements. It has been known since the work of Sherrington that artificially stimulating appropriate regions of the brain and spinal cord can be sufficient to produce the twitches, scratches, blinks, and so forth that characterize a number of reflexes, including conditionable behaviors (e.g., Mauk and Thompson, 1984). These regions could in principle serve as the required M's, at least for simple behaviors.

The question then becomes what role C's and M's might have in learning. At present, neurobiologists can tell us very little about how operant conditioning occurs (though for promising results at the invertebrate level see Hawkins et al. 1985; Cook and Carew 1988). A great deal more is known about Pavlovian or classical conditioning. It is therefore tempting to ask whether results in this area run against the grain of Dretske's semantic indicator. Of course, we first must show that there is sufficient formal similarity—in light of Dretske's theory—between operant and classical conditioning.

The central difference between operant and classical conditioning lies in what governs reinforcement during training. In operant conditioning, reinforcement is contingent on something the animal does. In our earlier example, the reinforcing event (food) is withheld until the rat treads on a lever. In a classical conditioning experiment, reinforcement is completely independent of anything the animal does; it is linked instead to environmental cues or stimuli. Thus, in Pavlov's experiment, the reinforcer (meat powder) was contingent on the ringing of a bell but independent of the dog's movements.

This contrast is shown diagrammatically in panels A and B of Figure 2, where training contingencies introduced by the experimenter are shown as single arrows. Within each panel, the diagram on the left shows the potential connections that exist prior to conditioning (before). The diagrams on the right show the conditioned state (after). Notice that in operant conditioning, the reinforcing event R is linked to a movement M3, while in classical conditioning it is linked to a stimulus F3. A second difference is in the range of conditionable responses. The particular reinforcer in a traditional classical conditioning experiments restricts the conditioned behavior to movements produced by the reinforcer itself. Thus Figure 2B contains a single M, while in 2A, many M's could be connected to particular C's.
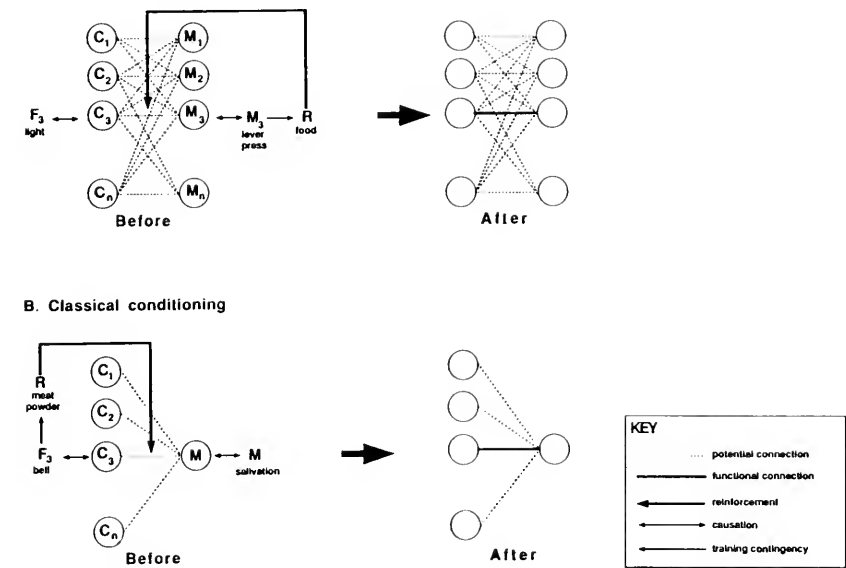
FIGURE 2. Contrasts between operant and classical conditioning. In operant conditioning (A), the reinforcing event R is contingent on a particular movement M3, but R could in principle have been contingent on any of the range of movements M1 to Mn, thus increasing the number of conditionable responses. In classical conditioning (B), R is independent of M and only the M produced by R becomes conditioned. F, C, and M are as in Figure 1; symbols are given in the key.

Returning briefly to the themes of the first section, let's ask what is special about operant conditioning, such that mental representation gets a foothold here. There are two elements to the answer. First, we have seen that for representation to be involved, "structuring" events must be among those in need of explanation. In the thermostat example, it was only of interest to ask why this sensor was wired to the garage door because everyone knows that insofar as the laws of physics are concerned, it could have been attached to any electrical appliance, limited only by the electrician's caprice. The situation is similar for operant conditioning. In the previous example the rat comes to lever-press (M) in the presence of the light (indicated by C), but we know that we could have conditioned him to a different light, or a tone, or a buzzer and so on. Similarly, lever-pressing is not the only conditioned response that could have become associated with the light. The second element of the answer is that indicator properties of C must be an important part of the explanation of the structuring event. When we inquire why *this* C became attached to *this* M (and not some other C), part of the answer must be that only this C indicated the relevant aspects of the environment. In the operant conditioning example, C3 indicates the light F3. C3 becomes associated with lever-pressing because of what the *light*

happens to be correlated with. As this particular operant experiment is conducted, the experimenter's M3–R contingency is only in force when the light is on. Only light was coupled to the M3–R contingency, therefore only the indicator of the light, C3, gained control over M3.

Despite their differences, both the operant and classical treatments are equally well viewed as cases where structuring explanations are required. One can just as easily ask why *this* C becomes connected to *this* M (and not some other C to some other M) in classical conditioning as in operant conditioning. This is because in classical conditioning too there are many potential C–M links, as Figure 2 suggests. Nor should the differences between the experimental contingencies underlying operant and classical conditioning obscure the fact that in both cases, the C that becomes associated with M is the one that indicates the relevant environmental event F, and that the relevant F is determined by the same considerations in the two cases. The relevant F is always the stimulus that is correlated with reinforcement. The fact that in one case (classical conditioning) F is correlated with the reinforcing event itself, while in the other (operant conditioning) F is correlated with the contingency of two other events (M and R) does not affect the character of the structuring explanation given. It is still which F a C indicates that determines its associability. Thus, there is an important formal similarity between operant and classical conditioning. Each is an instance of selective connection of C's to M's, and selectivity is in both cases achieved because only one C has the right indicator role.

We are now in a position to ask how well what is known about classical conditioning comports with the Dretskean models of Figure 2. The system in which the neural mechanism of classical conditioning has been pursued most thoroughly is the gill withdrawal reflex of the marine mollusk *Aplysia*. The *Aplysia* nervous system is simple by human standards yet this animal, or its close phylogenetic relations, performs favorably under a wide range of complex conditioning procedures including second-order conditioning, blocking, operant conditioning, food aversion training, and conditioned emotional response (for reviews, see Byrne 1987; Carew and Sahley 1986).

In response to a moderate tactile stimulus to a fleshy spout called the siphon, the animal withdraws its gill apparatus into a cavity on its back. This is accomplished in large measure by the activity of siphon sensory neurons that synapse directly on motor neurons that produce withdrawal of the gill. Shocking the animal's tail causes a much stronger gill withdrawal and can be used as a reinforcing stimulus. When activity of a siphon sensory neuron is paired with tail shock in a classical conditioning experiment, gill responses are enhanced to that sensory neuron but not another siphon sensory neuron whose activity was not paired with tail shock (Walters and Byrne 1983; Hawkins et al. 1983). Physiolog-

ical studies have subsequently localized the site of classical conditioning to the sensory neuron to motor neuron synapse.

A model for how classical conditioning occurs in *Aplysia* is shown in Figure 3. For simplicity, only two of the more than 20 sensory neurons are shown. Each sensory neuron (C3 and C2) can be considered an explicit indicator of touch (F1 and F2) to a particular region of the siphon. Reinforcement is subserved by a facilitating interneuron that responds to tail shock. When the experimenter introduces a specific contingency between activity in sensory neuron 1 and tail shock, the connection between this neuron and the motor neuron (M) is selectively strengthened, as the diagram on the right illustrates using the size of triangles to represent relative synaptic strength. The striking formal similarities to Dretske's mode of selective recruitment are clear from comparison of Figures 2B and 3. In each case a reinforcing mechanism acts to enhance just one of the possible C–M connections. In particular, the C–M connection is strengthened only for the indicator of the stimulus that in turn predicted the reinforcing event.
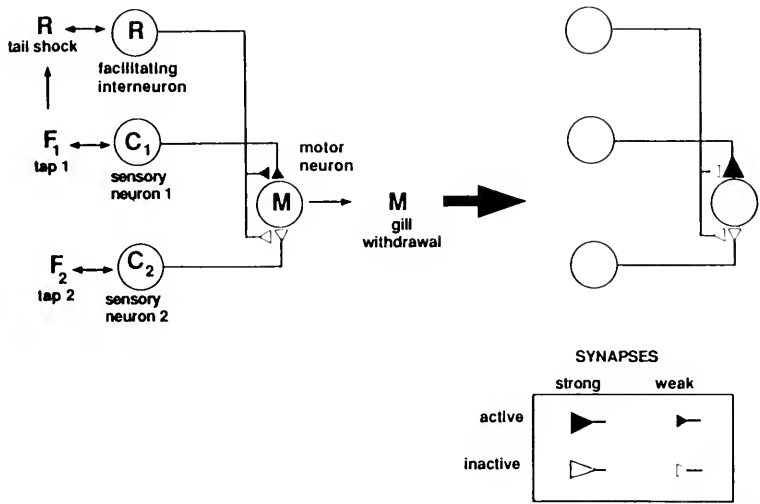
FIGURE 3. Model of classical conditioning in the gill withdrawal reflex in *Aplysia*. Circles represent the indicated sensory, motor, and facilitating interneurons contributing to the reflex. Triangles represent connections (synapses) between neurons. As shown in the key, active synapses are black and the size of a triangle indicates the strength of the connection. Each sensory neuron responds to stimulation (tap) of a different location on the skin. When stimulation of a particular location (e.g., location 1) is reinforced by tail shock, the connection from the corresponding sensory neuron to the motor neuron producing gill withdrawal is specifically enhanced.

Let us return to the example of the garage door wired to the thermostat, now with an eye for physical mechanism. Initially, of course, the thermostat and door were wholly unconnected. The connection was

established by the electrician who strung the wires from one to the other. One might call this the *ex nihlo* model of learning since the components start off with no connection of any kind. The consensus that is emerging from studies of how learning actually occurs points to a somewhat different model (Byrne 1987). We have already seen in *Aplysia* that learning consists of selectively enhancing existing connections. This also seems to be the case in less completely understood reflexes from a wide range of species, invertebrate and vertebrate alike. It is as though the wires are already in place but the contacts weak. This would be bad news for Dretske if his account of semantics required the *ex nihlo* model as his thermostat example would suggest. But there is no reason to suppose this is the case. As we have seen, for a particular C to count as a mental representation, all that is required is that its indicator properties be an important part of the structuring explanation. In particular, it is enough to show that the selectivity inherent in the structuring event (why this C and not *that* one) is attributable to the indicator properties of the C that wins out. The details of the resulting physical events that eventuate in the required connectivity causation are immaterial, so long as the causal role of C is essential to explaining what initiated them. It therefore does not matter whether, because of its particular capacities for indication, a certain C "grows" an entirely new connection or undergoes the strengthening of one already in place.

It is too early of course to say precisely how general the neural mechanisms of conditioning will turn out to be. But given the trend toward conservation during evolution, there is at least some reason to expect that the mechanics of classical conditioning will be fundamentally the same in other systems. It is interesting in this regard that the basic principles of simple classical conditioning can be used, at least theoretically, to construct models of more complex forms of conditioning, including operant conditioning. This means that the theme of selective recruitment operating on explicit indicators of sensory events may prove quite general indeed.

PROTO-BELIEFS AND INTENTIONALITY

Dretske, as we have seen, takes the case of operant conditioning as a simplified model for the sort of mental representation to be found in conscious human beliefs. The "proto-beliefs" that rats or sea slugs acquire in operant conditioning are much the same, Dretske maintains, as Dretske's own belief that there is a beer left in the fridge. The only major difference Dretske sees between the proto-belief of the sea slug and the full-fledged belief of the human is that humans have more of

them. The beliefs of adult humans are embedded in much richer networks. However, what we propose to argue in this section is that Dretske's "proto-beliefs" are actually very poor models for the sorts of beliefs presupposed by common-sense psychological explanations of human action.

At the core of our argument is the fact that adult human belief, as conceived of by commonsense psychology and as exploited in reason-giving explanations of human behavior, is capable of being very specific in the way it represents the world. What Fred believes is that there is a bottle of beer in the fridge, not that there is a bottle of ale, or stout. Oedepus believed that Jocasta would be a good person for him to marry. He did not believe that his mother would be a good person for him to marry. It is even possible for a person to believe that p and not believe that q, despite the fact that p and q are logically equivalent. The standard examples here involve propositions whose logical equivalence is not obvious, and not known to the person in question. Another facet of the precision that is possible in the common sense concept of belief is that people can have beliefs about things that they cannot reliably identify. Fred may believe that the gem in his wife's wedding ring is a diamond, though he has no idea how to distinguish real diamonds from fake ones. It is our contention that Dretske's notions of mental representation and proto-belief are incapable of attaining this sort of fine grained discrimination. Thus Dretske's constructs will not do as models for full-blown human intentionality.

At the heart of the problem is the fact that indication is a promiscuous relation. An internal state that indicates one external state of affairs will typically indicate many others as well—recall the example of the tree ring. As an indicator state, a C in Dretske's schematic formulation, comes to be a full-fledged representation of some state of affairs, F (and a proto-belief that F is the case), when it acquires the function of indicating F. But the same indicator will typically also indicate various other states of affairs: G, H, etc. How are we to determine which of the various states of affairs that C indicates it comes to represent? As we have seen, Dretske's answer is that C represents the state of affairs it has the function of indicating. However, it is our contention that the notion of function is simply not sufficiently discriminating to distinguish between the various state of affairs that C indicates.

All this will be a bit clearer if we consider a specific example. We'll focus on one that Dretske introduces himself. Monarch butterflies store a noxious substance from the milkweed plants on which they feed. When a bird eats a monarch butterfly it finds it foul tasting, and quickly learns to avoid monarchs in the future. Exploiting this fact, another species of butterfly, the viceroy, has evolved to resemble the monarch. However, the

viceroy has not evolved the monarch's system of storing a noxious substance. A bird that eats a viceroy is not punished by a foul taste. The viceroy is a mimic, an evolutionary freeloader, it gets by on looks alone.

Now consider the situation of the bird that has eaten a number of monarch butterflies and learned to avoid them. We may suppose that it has also eaten a viceroy or two, with no ill effects. Let's also suppose that this bird has encountered and eaten a monarch which, for one reason or another had not stored any noxious substance, once again with no ill effects. Now, if Dretske is right, there is some internal state—some C—in the bird that is lawfully correlated with monarch butterflies. Thus C indicates monarchs. However, C also indicates a larger class: the class consisting of monarchs and viceroys. The bird avoids viceroys because viceroys trigger state C. C also indicates a smaller class: the class of noxious monarchs; it was members of this class that provided the positive punishment that led to C being made a cause of avoidance behavior. When the bird has been conditioned—when C has come to cause avoidance behavior that it did not previously cause—what does C represent? Does it represent monarchs? or just noxious monarch? or monarchs and viceroys? Dretske answers this sort of question as follows:

> C will normally indicate a great many things other than F. Its indication of F is, therefore only "one component" of its natural meaning.[1] Nonetheless, it is this single component that is promoted to representational status . . . because it is C's indication of F, not its indication of (say) G or H, that explains it causing M. Hence, it becomes C's function to indicate F, not G or H (p. 84).

Making the substitutions relevant to the present example we obtain:

> C will normally indicate a great many things other than *noxious monarchs*. Its indication of *noxious monarchs* is, therefore only "one component" of its natural meaning. Nonetheless, it is this single component that is promoted to representational status . . . because it is C's indication of *noxious monarchs*, not its indication of (say) *monarchs* or [the class] *monarchs and viceroys*, that explains its causing M. Hence, it becomes C's function to indicate *noxious monarchs*, not *monarchs* or [the class] *monarchs and viceroys*.

Thus we see Dretske's answer is that C must represent only the reinforceable class, the noxious monarchs. But to draw the line here runs against the grain of common sense. Consider the example of the rat undergoing operant conditioning of lever presses to the light. Suppose lever pressing is reinforced by food, still only when the light is on,

---

1. As Dretske makes clear elsewhere (p. 55), "to mean" in the sense of natural meaning is, for him, synonymous with "to indicate."

but now on a less than 100 percent schedule. Drawing the obvious analogies, the light flashes associated with reinforced lever presses are equivalent to the noxious monarchs, and so it is only this class that the relevant C represents. But surely any coherent theory of animal informa-tion processing ought to maintain, to the contrary, that what C repres-ents, and what the animal learns about, is the light—any and all of its flashes, not just the subclass of its reinforcement associated occur-rences. Without this, there is no account of the animal's persistence in lever pressing when the light is on, yet lever pressing is not reinforced.

Another reason it seems more natural to say that C represents all the light flashes is that the C triggering mechanism (whatever it is) has no means of telling reinforcement associated flashes from flashes unre-lated to reinforcement. The experimenters can tell the difference since obviously it is within their powers to inspect the state of the Skinner box on each trial, but from the point of view of the C in question, one flash is like any other. Of course, one monarch is like any other too, and here enters a new problem. From the point of view of the relevant C, one *monarch* is like any *viceroy.* This means that C represents monarchs—both kinds—and viceroys. The problem is that there appears to be no stopping. All things that trigger C (in a context where enough of them are associated with reinforcement to support conditioning) are represented by it. This means that C never misrepresents—the bird never has false beliefs. But the possibility of believing falsely is a property of human psychology just as essential as fine-grainedness of belief. On neither reading of representation do we find a suitable model of two truly central aspects of human mental representation.

It appears then that the notion of function is not strong enough to provide a well motivated way of deciding which of several quite different alternatives C represents. And if that's right, then Dretske's notion of representation and proto-belief will not be sufficiently fine grained to serve as the foundation for anything much like the commonsense notion of belief. If Fred believes that the insect he is looking at is a monarch, that belief is very different from the belief that the insect he is looking at is either a monarch or a viceroy. Indeed, the latter belief could well be true when the former is false.

This completes our case for one of the two claims we promised to illustrate in our introductory remarks. Along with Staddon (1988), we suspect that the study of intelligent and adaptive behavior in animals has provided relatively little insight into human psychology or human intelligence. Dretske's analysis of representation and proto-belief seems to us to be a case in point. His goal was to shed light on the structure of the intentional concepts we use in common-sense reason-giving expla-nations, and to explicate the strategy invoked in such explanation. But if we are right, then Dretske's efforts have only partially succeeded. The animal model has thrown light on beliefs of only the crudest sort and is

destined to blur the fine distinctions that make human beliefs truly human.

On the positive side, however, it seems to us that Dretske has at least made an intriguing beginning at explicating a strategy of explanation that is of importance in understanding adaptive behavior in animals. A complete neurobiological account detailing each step in the causal process from stimulus to behavior will not explain everything that needs explaining. It will not tell us why one rat presses the lever when the light goes on and its identical twin does not. It will not tell us why one hungry bird avoids viceroy butterflies while a conspecific does not. Explaining these facts requires more than circuitry; it requires an appeal to the history of the organism and the environment in which that history unfolded. The complete story about why certain birds avoid viceroy butterflies has two parts: First, they have an internal state which, in their environment, indicates both noxious monarchs and harmless viceroys. Second, there is a neural mechanism, the details of which are gradually becoming clear, which results in the indicator triggering avoidance behavior in just such circumstances. Plainly there is something quasi-intentional about such explanations. They involve both the internal workings of the organism and correlations with the environment. It is less clear whether any more richly intentional notion will play a role in the explanation of adaptive behavior in animals—whether notions like representation, misrepresentation and belief which play a central role in human reason-giving psychology have any work to do in the explanation of animal behavior. In our view the answer to this question can't be settled by a priori speculation. Only careful research and theory building will do. Should it turn out that the answer is negative, however, we won't be surprised or disappointed. Understanding adaptive behavior in animals is a profoundly interesting project even if it does not provide a useful model for purposive behavior in humans.

ACKNOWLEDGMENT

REFERENCES

Armstrong, D. M. 1973. *Belief, truth and knowledge.* (CUP, Cambridge).
Byrne, J. H. 1987. Cellular analysis of associative learning. *Physiological Reviews* 67:329–439.
Carew, T. J. and Sahley, C. L. 1986 Invertebrate learning and memory: from behavior to molecules. *Ann. Rev. Neurosci 9,* 435–487.

Cook, D. G. and Carew, T. J. 1988. Operant conditioning of identified neck muscles and individual motor neurons in *Aplysia. Abstra. Soc. Neurosci 14*, 607.

Dennett, D. C. 1987. *The intentional stance, Ch. 8 (MIT, Cambridge)*.

Dretske, F. 1988 *Explaining behavior* (MIT, Cambridge).

Hawkins, R. D., Abrams, T. W., Carew, T. J, and Kandel, E. R. 1983. A cellular mechanism of classical conditioning in *Aplysia:* activity-dependent amplification of presynaptic facilitation. *Science* 219, 400-405.

Hawkins, R. D., Clark, G. A., Kandel, E. R. 1985. Operant conditioning and differential classical conditioning of gill withdrawal in *Aplysia. Abstr. Soc. Neurosci. 11*, 796.

Hubel, D. H., and Wiesel, T. N. 1977. Functional architecture of macaque visual cortex. *Proc. Roy. Soc. (Lond) Series B 198,* 1-59.

Mauk, M.D. and Thompson, R. F. 1984. Classical conditioning using the inferior olive as the unconditioned stimulus. *Abstr. Soc. Neurosci. 10*, 122.

Mountcastle, V. 1957. Modality and topographic properties of single neurons of the cat's somatic sensory cortex. *J. Neurophysiology 20*, 408-431.

Ramsey, F. P. 1931. *The foundations of mathematics, and other logical essays* (RKP, London).

Staddon, J. E. R. 1988. Animal psychology: the tyranny of anthropocentrism. In P. P. G. Bateson and P. H. Klopfer (eds) *Perspectives in Ethology.* Vol. 8: Whither Ethology Plenum, London.

Walters, E. T., Byrne, J. H. 1983. Associate conditioning of single sensory neurons suggests a cellular mechanism for learning *Science 219*, 405-408.

ADAPTATION VERSUS PHYLOGENY: THE ROLE OF ANIMAL PSYCHOLOGY IN THE STUDY OF HUMAN BEHAVIOR

John Tooby and Leda Cosmides
Stanford University

ABSTRACT: Advocates of Darwinian approaches to the study of behavior are divided over what an evolutionary perspective is thought to entail. Some take "evolution-mindedness" to mean "phylogeny-mindedness," whereas others take it to mean "adaptation-mindedness." Historically, comparative psychology began as the search for mental continuities between humans and other animals: a phylogenetic approach. Independently, ethologists and now behavioral ecologists have placed far more emphasis on the niche-differentiated mental abilities unique to the species being investigated: an adaptive approach. We argue that the output of complex, dynamical systems can be dramatically changed by only minor changes in internal structure. Because selection acts on the consequences of behavior, the behavioral output of the psyche will be easily shaped by adaptive demands over evolutionary time, even though the modification of the neurophysiological substrate necessary to create such adaptive changes may be minor. Thus, adaptation-mindedness will be most illuminating in the study of cognition and behavior, whereas phylogeny-mindedness will be most illuminating in the study of their neurophysiological substrates. Similarly, a phylogenetic approach to cognition and behavior is likely to cause one to overlook our most interesting, complexly designed species-typical traits, whereas using animal psychology to exfoliate general principles of behavioral ecology represents our best hope of understanding humanity's many zoologically unique characteristics.

Darwin, with the publication of the *Origin of Species* (1859) and the *Descent of Man* (1871), united the human and the animal worlds into a single system by proposing an explanation for species' characteristics, including their similarities and differences from each other, in terms of the operation of intelligible natural causal processes. By tying all animals together in a single tree of descent, Darwin made the study of every species relevant to the study of every other species. Animals drawn from different species are separated only by phylogenetic distance; character differences separating different phylogenetic groups were produced either by chance, or they were driven by niche-differentiating selection pressures.

This scientific account of the nature of living things embedded human characteristics in the world of natural cause and effect, and constituted a radical attack on the nearly universally accepted doctrine of total human singularity: the idea of humans as a special divine creation, or as outside of nature, or as subject to principles—supernatural or otherwise—totally alien to those operating in the rest of the world. Darwin's departure was so radical that the human sciences are still trying to come to terms with it, and the debate about human singularity remains with us today. The guise in which this debate continually re-emerges changes—from claims about reason, or intelligence, or language, or learning, or emergent social processes, or the superorganic nature of human culture—but the attachment to the idea of humans as subject to entirely unique principles is, to this day, the centerpiece of persisting, anti-Darwinian arguments (e.g., Durkheim, 1962; Kroeber, 1952; Sahlins, 1976). Implicit in the Darwinian revolution is the recognition that however interesting and endearing we humans might be to ourselves, we are simply one species out of an entire ocean of species; if humans are not the product of unique principles, if we are simply one causal outcome in a larger scientific landscape, then there is potentially a general and principled science that encompasses the entire animal world (Darwin, 1871; Williams, 1966; Staddon, 1988; Tooby & De Vore, 1987).

There are three positions one might take on human singularity. The first is that humans are not unique, but rather are typical animals (e.g., Skinner, 1938, 1953, 1957; see also Macphail, 1987). Although this view is clearly true when humans and nonhuman animals are compared along certain dimensions, such as in terms of their biochemical and cellular processes, for many other dimensions such a view is difficult to sustain, given human agriculture, machine tools, VCRs, ping pong, nonsense verse, deconstructivist architecture, and so on. Much of manifest human behavior appears to be qualitatively different from that of other animal species.

A second possible position derives from traditional religious and philosophical views, echoed in modern form throughout the social sciences (e.g., Descartes, 1977; Durkheim, 1962; Evans-Pritchard, 1954; Geertz, 1973; Kroeber, 1952; Radcliffe-Brown, 1977; Sahlins, 1976): that humans are so singular, so incomparable, that they are either outside of nature, or at least a natural phenomenon *sui generis*, governed by their own special and incommensurate laws. More precisely, this position is that humans may (or may not) be legitimate objects of scientific study, but that the principles that created them or that govern them are unique, not derived from or applicable to any other species—that humans cannot be analyzed with conceptual tools other than those specially devised for understanding them. This, of course, means that that evolution is irrelevant, and that other animals are irrelevant, separated from humans by an unbridgeable gulf (Sahlins, 1976).

The third view is that humans have many unique, zoologically unprecedented properties that make us unlike any other species, but that this is not because humans are the product of unique principles particular to humans, but rather because we are the product of a unique combination of general evolutionary principles, which act across the field of animal life. Other forms of life also manifest zoologically strange features, such as the eusocial insects, with their separate castes linked to their bizarre genetic systems (Wilson, 1971). Yet, understanding them proved generally illuminating to our understanding of evolutionary principles, such as kin selection, that apply to all animals (Alexander, 1974; Hamilton, 1964; Williams & Williams, 1957; Williams, 1966; Wilson, 1975). By looking at each species as a unique combination of general evolutionary principles, it is possible to deduce what these natural causal principles are, and to see how, in each instance, they fit together in a unique, yet fully comprehensible way (Alexander, 1971, 1974; Tooby & DeVore, 1987).

Darwin himself, along with his contemporaries, realized that the most controversial of his claims was that the evolutionary perspective applied with equal force to the psychological as well as the physiological (Darwin, 1873). Since Descartes, educated belief was quite willing to hold that the physical body was a machine, subject to physical law, and that animals were automatons, like the water-powered robots in the gardens of Louis XIV. It was mental phenomena, believed to exist only in humans, such as reason, emotions, goal-seeking, language, and culture, that were separated off by religious belief and Cartesian dualism into the extra-physical, extranatural domain of the soul, the mental, the psychical (Descartes, 1977). It was this dualism that Darwin's much misunderstood concept of mental continuity was addressed to, and indeed was intended to refute: the dualistic claim that mental phenomena in humans represented a qualitatively different essence, a spiritual agency, constituting an abyss that evolutionary explanations could not bridge (Darwin, 1873). Darwin argued that mental faculties were explicable in the same evolutionary terms that accounted for the origin of species and the acquisition of their physiological characteristics—a position that Alfred Russell Wallace, co-originator of the theory of natural selection, was not willing to endorse: Wallace (1904) felt that human mental faculties required supernatural explanation. [Darwin's doctrine of mental continuity was subsequently overliteralized by early comparative psychologists into the idea that all animals existed on a linear continuum, with only quantitative differences in their capacities (see Hodos & Campbell, 1969, and Lockard, 1971, for a critical discussion of this issue; see Macphail, 1987, for a post-behaviorist defense of the doctrine of mental continuity)]. So Darwin's achievement was more than just the principled unification of the human, animal, and plant worlds: It was

also the principled unification of the mental and physical worlds, joining the mental and physical characteristics of humans and other animals into the same system of causation, created by mechanistic evolutionary principles. This allowed a radical transformation of biology and psychology as sciences (Dawkins, 1976; 1986). Instead of being descriptive, particularistic sciences, fueled by unguided observation, the cataloguing of phenomena, and the inductive, atheoretical search for regularities, biology and psychology could be seen as grounded in an elegant set of mechanistic evolutionary principles that provided a causal explanation of how each species acquired its distinctive characteristics—its design.

Since Darwin's time, despite the appealing prospect of a powerful, general, evolutionary psychology, evolutionary approaches to psychology have waxed and waned in popularity several times, with, for example, the long behaviorist interregnum, motivated by the anti-Darwinian belief that learning and environmentalist influences somehow insulate behavior from evolutionary shaping and analysis (Skinner, 1953; Boakes, 1984). As evolutionary-oriented psychologists ourselves, we hope that finally, evolutionary psychology has arrived on the scene permanently, anchored this time by a deeper and more balanced understanding of the nature-nurture issue (Daly & Wilson, 1983; Lehrman, 1970; Lorenz, 1965; Marler and Hamilton, 1966; Mayr, 1974; Staddon, 1983; Symons, 1987; Tinbergen, 1951; for discussion, see Cosmides & Tooby, 1987; Tooby, 1985; Tooby & De Vore, 1987; Tooby & Cosmides, in press), and by more comprehensive and useful models of the evolutionary process (e.g., Hamilton, 1964; Williams, 1966; Dawkins, 1976, 1982).

Unfortunately, however, there are fundamental differences in what an evolutionary perspective is thought to entail, generating substantial confusion. Two of the most important evolutionary principles accounting for the characteristics of animals are (1) common descent, and (2) adaptation driven by natural selection. Some take "evolution-mindedness" (Symons, 1987) to mean "phylogeny-mindedness," the search for phylogenetic continuities implied by the inheritance of homologous features from common ancestors. Others take evolution-mindedness to mean "adaptation-mindedness," the search for adaptive design, which usually entails the examination of niche-differentiated mental abilities unique to the species being investigated. Historically, comparative psychology began as the search for mental continuities between humans and other animals (Lockard, 1971; Hodos & Campbell, 1969): a phylogenetic approach that persists, in some measure, to the present (e.g., Boakes, 1984; Macphail, 1987). Independently, ethologists (subsequently joined by behavioral ecologists and many modern comparative psychologists) have placed far more emphasis on animal psychology as case studies in adaptive design (e.g., Alcock, 1989; Daly & Wilson, 1983; Dewsbury, 1984; Gould, 1982; Krebs & Davies, 1984; Staddon, 1988; Tinbergen, 1951; see discussion in Tooby & DeVore, 1987). The phylogeny-

minded tend to believe that the study of animal psychology provides observations that directly parallel human mental processes (Macphail, 1987), whereas the adaptation-minded tend to believe that the psychology of each species is distinct, and that direct leaps from one species to another are speculative and unfounded.

One can see the interplay of these sentiments in the ape-language controversy: Half of the scientific audience cheers for the apes, hoping they can duplicate human linguistic feats, while the other half is confident that the apes' linguistic abilities will prove very limited. The phylogeny-minded form the apes' cheering section: They reason that if a human can learn a language, then our nearest relatives should be able to do so as well. The adaptation-minded are skeptics in the ape language controversy: They (correctly) see the acquisition of a language as a species-specific mental ability, requiring highly complex and specialized cognitive mechanisms that are not likely to be shared by other primates, who were not selected to participate in communication through linguistic behavior (Chomsky, 1975).

Animals from different species are similar to each other in psychological architecture because of (1) common inheritance, (2) the same selection pressures operating on different species, or (3) both. Animals from different species differ in psychological architecture because of (1) independent descent, (2) the operation of different selection pressures on different species, or (3) random divergence. Both adaptive and phylogenetic components of the evolutionary approach have value, but their relative validity depends on exactly what level of psychological investigation they are applied to.

To understand why this is so, it is necessary to deal with the issue of the complexity and domain-specificity of psychological mechanisms. As we and others have argued elsewhere (Cosmides & Tooby, 1987; Rozin, 1976; Symons, 1987), an evolutionary perspective leads to the conclusion that although the psyche has some domain-general mechanisms, it must also include many domain-specific, function-specific mechanisms. This view draws support from artificial intelligence, whose history has largely been the history of discovering that information-processing procedures must be very complex indeed if they are to perform even very simple tasks (e.g., moving around a half a dozen blocks in a small area; see, e.g., Boden, 1977; Brown, 1987; Minsky, 1986; Cosmides & Tooby, 1989). Work in cognitive science and artificial intelligence (AI) has shown that mechanisms capable of solving even supposedly simple real-world cognitive tasks must contain very complex "innate" prespecified procedures and/ or information, matched narrowly to the structural features of the domains within which they are designed to operate (Marr, 1982; Chomsky, 1975, 1980; Fodor, 1983). AI programs are complex and function-specific because the world is itself complex in ways that are not logically analyzable or deducible without an enormous amount of *a*

priori knowledge: In order to solve a task, you must already know a great deal about the nature of the circumstances in which the task is embedded. Of course, for "natural intelligence," as opposed to artificial intelligence, the origin of such necessary *a priori* knowledge and procedures is readily understood: Hundreds of thousands of generations of exposure to recurrent adaptive problems selected for minds that came preequipped with mechanisms tuned to solving those problems (Cosmides & Tooby, 1987).

The process of adaptation through natural selection is what constructs mental mechanisms, with their functional characteristics; common descent can only conserve them. Adaptations evolve in response to selection pressures, some of which are very general (such as the requirement to see, or to function in three-dimensional space), whereas others can be ranked as increasingly specific, down to those selection pressures that act uniquely on a single species. Adaptations, consequently, will range from those that many species hold in common, to those held by some, to those specific to a single species. For these reasons, the adaptive approach often, but by no means always, involves claims of niche-differentiated mental abilities unique to the species being investigated. Shepard's (1984) work on the mental representation of motion in space, or Staddon's (1987) work on learning as inference, for example, represent the analysis of universal selection pressures general to the psychology of different animal species. Language acquisition, production, and comprehension, on the other hand, are the result of selection pressures that have only impinged on humans, although the function language serves for humans may reflect selection pressures that are broader in scope. This is why adaptation-minded evolutionists are so skeptical of expansive claims about the linguistic abilities of apes: Language was simply not part of their social environment, and clearly requires complex, specialized, innate psychological mechanisms, as Chomsky (1975, 1980) has so persuasively argued.

If a species is analyzed at the level of total behavioral output, and viewed as an integrated system, then species look very different from one another, and reasoning from species to species appears conjectural. For example, the early literature on territoriality or aggression (Ardrey, 1966; Lorenz, 1966[1]), which reasoned from the behavior of a scattered handful of species to human violence and warfare, had little to recommend it. Similarly, anthropological models of human evolution have been led seriously astray by the series of animal models that have been proposed: baboons, or chimpanzees, or hunting dogs behave in a certain fashion, and therefore it was asserted that early humans did as well (Tooby & DeVore, 1987).

---

[1]In fairness, it should be pointed out that this popular book (Lorenz, 1966) was very uncharacteristic of Lorenz's other work, in which he carefully explored and documented species-specific adaptation.

However, behavioral output is the combined simultaneous output of all psychological mechanisms activated in an animal at the time: A change in any one mechanism may affect the combined interacting output, making this level the most variable between species. Consequently, the only kind of analysis that makes this kind of approach valid is one that matches different species' behavior against selection pressures, finding regularities in their relationships, as in behavioral ecology (see, e.g., Clutton-Brock & Harvey, 1979). Behavioral ecology is the study of selection pressures, and of the behaviorally-described adaptations that have evolved in response to them.

However, if one drops below the level of total behavioral output, and as one decomposes the psyche into different psychological mechanisms, continuities (and discontinuities) from one species to another become far more apparent. Substantial design complexity is required to solve most specific adaptive problems, and this necessary complexity has several implications for whether the study of animal behavior is relevant to the study of human behavior:

1. Because selection tends to work step by step, complex systems take time to evolve, and wholly new psychological mechanisms do not appear rapidly. The amount of complex psychological adaptation to a problem will be proportional to the length of time selection has been operating, as well as the intensity of the selection pressure. For example, mother–infant relations are considerably more ancient than father–infant relations, and so mothers can be expected to have more specialized and reliable psychological mechanisms than fathers (Tooby & DeVore, 1987). Many human maternal psychological mechanisms may be shared among apes, or monkeys, or even mammals (e.g., oxytocin release during birth triggering imprinting of the mother on the neonates; Konner, 1982). On the other hand, significant paternal care is a human evolutionary novelty within the great ape lineage, and hence is an adaptively-driven, niche-differentiated selection pressure operating to produce psychological mechanisms unique to human males. However, paternal care, though recent in the sense of having emerged after the human lineage split off from the other great apes, may still be several million years old. Language is another case of a post-divergence psychological capacity. Very recent developments, such as agriculture, cities, or contraception, can be expected to have called forth no new psychological mechanisms or complex adaptations, because they have not existed long enough for selection to act (Symons, 1987; Cosmides & Tooby, 1987).

2. The requirement of complex design for effective psychological mechanisms means that, once a problem has been solved in an evolutionary lineage, the psychological mechanism generally will be retained, passed on to all descendent lineages. Just as with physical organs, mental organs (Chomsky, 1980) can show substantial structural continuity from species to species. This is what makes psychological experi-

mentation on one species valuable and illuminating for related species. Such continuity from species to related species will be present to the extent that the mechanism evolved prior to their phylogenetic divergence, and to the extent that the selection pressures responsible remain constant from species to species. Over time, mechanisms may indeed be modified or refined, but it is relatively unlikely that such mechanisms will be rapidly or completely supplanted. This has particular application to humans, where an entire range of standard adaptive functions, like eating, parenting, mating, and so on, have been widely held to have been entirely supplanted by equipotential, culture-learning mechanisms (Geertz, 1973; Kroeber, 1952; Sahlins, 1976). Anthropologists frequently claim that although humans may behave in certain respects like closely related species, it is for entirely different, "cultural," reasons (Sahlins, 1976).

3. If psychological architecture involves an intricate series of very complex, domain-specific mechanisms, then adaptation-mindedness is more important if the level being examined is behavior or the cognitive structure of the mechanisms regulating behavior. Correspondingly, phylogeny-mindedness may be more important in examining the physiological and neurological substrates of behavior. This is because the output of complex, dynamical systems can be completely altered by only minor changes in internal structure (Sparrow, 1982). Where output of a system is determined by the combinatorial interaction of many subsystems, changing any part of the interaction can completely change the output behavior of the system: Turning off the ignition of the car, or turning the steering wheel, or changing the diameter of one wheel, or removing the distributor, does not incrementally alter the behavior of the car; it changes its behavior qualitatively and dramatically.

The thought experiments of Valentino Braitenberg (1984) elegantly demonstrate that the behavioral output of complex, dynamical systems analogous to psychological mechanisms can be radically modified by minor changes in internal structure. In his book, *Vehicles: Experiments in Synthetic Psychology*, Braitenberg systematically explores the properties of simple "vehicles" (organisms) that consist of sensors connected to motors, and shows the richly intricate behavioral complexity that can emerge from even very simple structures. Step by step he shows how the behavior of these simple vehicles can be radically refashioned by an incremental change in a motor's threshold of activation, by a shift in the relative sensitivity of sensors, by changing the function relating stimulus intensity to motor speed from monotonic to nonlinear, by adding internal connections, and so forth. The more sensors and internal connections the vehicles had, the more a minor change in a subcomponent would transform the final behavioral output of the system in complex and interesting ways.

This finding translates directly from engineered or imagined systems to biological ones. Small lesions or the administration of minute quantities of drugs may have striking effects on behavior. Embryological processes are notorious for their sensitivity (Gilbert, 1985; Gould, 1977). In fact, even in fully matured and developmentally stabilized organisms, it is difficult to find psychological mechanisms that to do not display this general characteristic. For example, entire constellations of behaviors can be turned on or off in rats by the relative balance of two neurotransmittors, norepinephrine and serotonin, a process that can be experimentally manipulated by a minor drug-induced change (Ellison, 1979). These experimental manipulations could be heritably duplicated by, for example, a shift in the thresholds at which these neurotransmitters act. According to Ellison, serotonin-depleted rats engage in behaviors that are typical of rats "out in the world"—feeding, fighting, watching for predators, and so on. In familiar environments, serotonin-depleted rats spend more time out of their burrows: they eat more food and take their meals more frequently, they are hyperactive and exploratory, and they become hyperaggressive and competitive, especially at feeding time. In novel environments, they appear "frightened and paranoid": they are hypervigilant, they "freeze" frequently, and they try to stay out of sight. In contrast, norepinephrine-depleted rats act somewhat like rats who are at rest, safe in their burrows. In a familiar environment they tend to stay in their burrows, and when they do venture out they are inactive, are the last to come to feedings, underconsume food, and tend to lose battles, gradually falling to the bottom of the dominance hierarchy. In novel environments they do not show the predator-wariness of the serotonin-depleted rats: they are not vigilant, they rarely "freeze," and they do not hide.

In this case, a chemical switch turns entire suites of behaviors on and off, modulating a (normal) animal's behavior in adaptively appropriate ways. Once such switches have evolved, however, relatively modest modifications in them can form the substrate for substantial subsequent evolutionary change. A case can be made that the evolutionary divergence of the spotted hyena *(Crocuta crocuta)* reflects such a process (the following discussion is based on Alcock, 1989: pp. 207–209; Ewer, 1973; Gould, 1982: pp. 463–466; Gould, 1981; Kruuk, 1972; Racey & Skinner, 1979; and Stewart, 1987).

Female spotted hyenas present a constellation of traits, both behavioral and morphological, that are unusual in the mammalian world. Unlike the typical mammalian pattern, and more importantly, unlike all other hyena species, female spotted hyenas are larger than the males, lead their clans in hunting and territorial defense, and tend to dominate the males, even those that are slightly larger than themselves. In addition, the female's clitoris and labia are enlarged and folded such that

they resemble the penis and scrotum of the male hyena. As a result, the two sexes are difficult to distinguish by appearance.

It appears that this entire constellation of traits, which represents a marked departure from the typical mammalian pattern, may have evolved in substantial measure via a modification of a simple chemical switch. Unlike other hyenas, which live in small groups and hunt small game, spotted hyenas live in large clans of up to 100 individuals, hunt large game, and compete fiercely for meat from the animals they kill. Stewart (1987) has argued that these factors have selected intensively for aggressive, dominant females. Because androgens, such as testosterone, regulate aggressivity and dominance in mammals, Ewer (1973) hypothesized that female spotted hyenas would have a high level of androgens in their blood. This proved to be the case: in the spotted hyena, the blood androgen level of females is equal to that of males (Racey & Skinner, 1979).

Selection for dominant, aggressive females via the modification of this chemical switch appears to have had an interesting side-effect: the development of a sham penis and scrotum in females. Testosterone is the agent responsible for masculinizing mammalian embryos: the same embryonic tissue will develop into a penis and scrotum, or into a clitoris and labia, depending on whether it is exposed to high levels of testostrone during the critical developmental period. Because a pregnant spotted hyena has (phylogenetically) unprecedented levels of androgens in her blood, the female fetuses she carries are bathed in androgens, and, as a consequence, develop external genitalia of a male, as well as an array of other phenotypic properties phylogenetically and developmentally more characteristic of males. In comparison, the adult females of other hyena species have low blood androgen levels, and their daughters do not develop male-like characteristics such as greater size, aggressiveness or sham penises and scrotums (Racey and Skinner, 1979). Once present, the ancestral female spotted hyena's sham genitalia were further elaborated by natural selection, and they are now used in the kind of communicative social displays characteristic of all male hyenas.

We are not arguing that all of the integrated adaptively organized traits that differentiate spotted hyenas from their closest relatives were brought about by a single change. Given the evolutionary time available to improve and elaborate adaptively organized systems of traits after speciation, and the improbability that a single change could yield exactly the outcome necessary, it would be very surprising if a single change were all that was involved. There are undoubtedly many additional modifications throughout the design of the spotted hyena's inherited phenotype that were tuned and elaborated through subsequent selection acting on its psychology and morphology. For example, the female mimicry of the male genitalia appears to be too perfect to have been brought about entirely by androgen induction. Nevertheless, there is substantial evi-

dence suggesting that the major avenue of change was a simple modification in the level of circulating androgens in adult females.

Thus, a minor quantitative change in the internal structure of ancestral spotted hyenas—a change in the level of blood androgens—may have been primarily responsible for a dramatic transformation in the behavior, the psychology, the social system, and even the morphology, of the female spotted hyena, differentiating it not only from its phylogenetically closest relatives, but from the mammalian baseline as well. One can see in this account that in terms of the social system and its associated behaviors, spotted hyenas seem drastically transformed from their phylogenetic baseline, both local (other hyenas) and global (other mammals). Yet, the more their design is decomposed into physiological subcomponents, the more continuity appears, until the essence of their divergence can be traced to a rather modest modification in their developmental physiology. If a single change at the physiological level can accomplish so much, what can many minor changes at the physiological level cumulatively accomplish at the behavioral level?

A system as complex as a mammalian psyche has many interacting parts, providing a wealth of potential changes, and a change in any one can change the final behavioral output (Tooby & Cosmides, in press). Because selection acts on the consequences of behavior, the behavioral output of the psyche will be easily shaped by adaptive demands over evolutionary time, even though the modification of the innate neurophysiology necessary to create such adaptive changes may be comparatively minor. As Ernst Mayr points out, even in tracking the truly massive aggregate changes from a reptilian ancestor to a bird or mammal descendent, "we are astonished at how few are truly new structures. Most differences are merely shifts in proportions, fusions, losses, secondary duplications, and similar changes that do not materially affect what the morphologist calls the plan of the particular type" (Mayr, 1976 pp. 96). The power of such modifications undercuts arguments about phylogenetic constraint (Gould & Lewontin, 1979) made at the adaptive and behavioral levels: The vast differences in social behavior between humans, chimpanzees, orangutans, and gorillas is eloquent testimony to how little phylogeny constrains the adaptive differentiation of psychologies among even closely related species. Because combinatorial interaction magnifies the impact of changes on output, the sculpting influence of selection will be strongly felt on behavior and on psychological mechanisms described in terms of functional design. Reciprocally, even striking behavioral differences between related species may be traced to relatively minor changes in underlying neurophysiological arrangement.

Thus, by specifying what "level" or kind of biological phenomena are involved, the controversy over the relative freedom of selection to create adaptive design versus the limiting role of phylogenetic constraint (Dawkins, 1982; Gould & Lewontin, 1979; Williams, 1966) may be clarified

(as may the otherwise well-explored issue of the relative role of homology and analogy in evolutionary analyses; Atz, 1970). We suggest that the nature of complex design makes the search for continuities among animals (including humans) helpful and illuminating at the neurophysiological level, where structural homologies are easily recognizable and readily traceable through related species. However, it is only adaptation-mindedness, and more specifically, the analysis of niche-differentiating species-specific selection pressures, that can illuminate the zoologically unique features of any species' psyche, including our own.

ACKNOWLEDGMENTS

REFERENCES

Alcock, J. (1989). *Animal behavior: An evolutionary approach. (fourth edition)*. Sunderland, MA: Sinauer.

Alexander, R. D. (1971). The search for an evolutionary philosophy of man. *Proceedings of the Royal Society of Victoria, 84*, 99-120.

Alexander, R. D. (1974). The evolution of social behavior. *Annual Review of Ecology & Systematics, 5*, 325-383.

Ardrey, R. (1966). *The territorial imperative*. New York: Atheneum.

Atz, J. W. (1970). The application of the idea of homology to behavior. In Aronson, L. R., Tobach, E., Lehrman, D. S. & Rosenblatt, J. S. (Eds.), *Development and evolution of behavior*, pp. 53-74   San Francisco: Freeman.

Boakes, R. A. (1984). *From Darwin to behaviorism: Psychology and the minds of animals*. London: Cambridge University Press.

Boden M. (1977). *Artificial intelligence and natural man*. New York: Basic Books.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Brown, F.M. (1987). *The frame problem in artificial intelligence*. Los Altos, CA: Morgan Kaufman.

Chomsky, N. (1975). *Reflections on language*. New York: Random House.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Clutton-Brock, T. H., & Harvey, P. H. (1979). Comparison and adaptation. *Proceedings of the Royal Society of London (B), 205*, 547-565.

Cosmides, L. & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In Dupre, J. (Ed), *The latest on the best: Essays on evolution and optimality*, pp. 227-306. Cambridge, MA: MIT Press.

Cosmides, L. & Tooby, J. (1989). Evolutionary psychology and the generation of culture, Part II. Case study: A computational theory of social exchange. *Ethology and Sociobiology, 10*, 51-97.

Daly, M. & Wilson, M. (1983). *Sex, evolution, and behavior. (second edition)*. Boston: Willard Grant Press.

Darwin, C. (1859). *On the Origin of Species*. London: Murray.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: Murray.

Darwin, C. (1873). *The expression of the emotions in man and animals*. New York: Appleton.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype: The gene as the unit of selection*. Oxford and San Francisco: W. H. Freeman.

Dawkins, R. (1986). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. New York: Norton.

Descartes, R. (1977). *The essential writings*. New York: Harper & Row.

Dewsbury, D. (1984). *Comparative psychology in the twentieth century*. New York: Van Nostrand Reinhold.

Durkheim, E. (1962). *The rules of the sociological method*. Glencoe, Illinois: Free Press.

Ellison, G. D. (1979). Chemical systems of the brain and evolution. In Oakley, D. A. & Plotkin, H. C. (Eds.), *Brain, behaviour and evolution*, pp. 78-98. New York: Methuen.

Evans-Pritchard, E. E. (1954). *Social anthropology*. Glencoe, Illinois: Free Press.

Ewer, R. F. (1973). *The carnivores*. Ithaca, NY: Cornell University Press.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Gilbert, S. F. (1985). *Developmental biology*. Sunderland, MA: Sinauer.

Gould, J. J. (1982). *Ethology: The mechanisms and evolution of behavior*. New York: Norton.

Gould, S. J. (1977). *Ontogeny and Phylogeny*. Cambridge, MA: Harvard University Press.

Gould, S. J. (1981). Hyena myths and realities. *Natural History, 90*, 16-24.

Gould, S. J. & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist program. *Proceedings of the Royal Society of London, 205*, 581-598.

Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology 7*, 1-52.

Hodos, W. & Campbell, C. B. G. (1969). Scala naturae: Why there is no theory in comparative psychology. *Psychological Review, 76*, 337-350.

Konner, M. (1982). *The tangled wing*. New York: Holt, Rinehart, & Winston.

Krebs, J. R. & Davies, N. B. (1984). *Behavioural ecology: An evolutionary approach. (second edition)*. Sunderland, MA: Sinauer.

Kroeber, A. L. (1952). *The nature of culture*. Chicago: University of Chicago Press.

Kruuk, H. (1972). *The spotted hyena*. Chicago: University of Chicago Press.

Lehrman, D. S. (1970). Semantic and conceptual issues in the nature-nurture problem. In Aronson, L. R., Tobach, E., Lehrman, D. S. & Rosenblatt, J. S. (Eds.), *Development and evolution of behavior*, pp. 17-52. San Francisco: W. H. Freeman.

Lockard, R. (1971). Reflections on the fall of comparative psychology: Is there a message for us all? *American Psychologist, 26*, 22-32.

Lorenz, K. (1965). *Evolution and the modification of behavior*. Chicago: University of Chicago Press.

Lorenz, K. (1966). *On aggression*. London: Methuen.

Macphail, E. (1987). The comparative psychology of intelligence. *Behavioral and Brain Sciences, 10*, 645-689.

Marler, P. & Hamilton, W. J. III. (1966). *Mechanisms of animal behavior*. New York: Wiley.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Mayr, E. (1974). Behavior programs and evolutionary strategies. *American Scientist, 62*, 650-659.

Mayr, E. (1976). The emergence of evolutionary novelties. In: *Evolution and the diversity of life*, pp. 88-113. Cambridge, MA: Harvard University Press.

Minsky, M. (1986). *The society of mind*. New York: Simon & Schuster.

Racey, P. A. & Skinner, J. D. (1979). Endocrine aspects of sexual mimicry in spotted hyenas *Crocuta crocuta*. *Journal of Zoology, London, 187*, 315-326.

Radcliffe-Brown, A. R. (1977). *The social anthropology of Radcliffe-Brown*. London: Routledge & Kegan Paul.

Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In Sprague, J. M., & Epstein, A. N. (Eds.), *Progress in psychobiology and physiological psychology*, pp. 245-280. New York: Academic Press.

Sahlins, M. (1976). *The use and abuse of biology: An anthropological critique of socio-biology*. Ann Arbor: University of Michigan Press.

Shepard, R. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91*, 417-447.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Skinner, B. F. (1957). *Verbal behavior.* New York: Appleton.

Sparrow, C. (1982). *The Lorenz equations, bifurcations, chaos, and strange attractors.* Berlin: Springer-Verlag.

Staddon, J. E. R. (1983). *Adaptive behavior and learning.* Cambridge, UK: Cambridge University Press.

Staddon, J. E. R. (1987). Learning as inference. In Bolles, R. C. & Beecher, M. D. (Eds), *Evolution and learning,* pp. 59-77. Hillsdale, NJ: Erlbaum.

Staddon, J. E. R. (1988). Animal psychology: The tyranny of anthropocentrism. In Bateson, P. P. G. & Klopfer, P. (Eds), *Perspectives in ethology. Vol. 8: Whither ethology?,* pp. 123-135. London: Plenum.

Stewart, K. J. (1987). Spotted hyenas: The importance of being dominant. *Trends in Ecology and Evolution, 2,* 88-89.

Symons, D. (1987). If we're all Darwinians, what's the fuss about? In Crawford, C. B., Smith, M. F., & Krebs, D. L. (Eds), *Sociobiology and psychology,* pp. 121-146. Hillsdale, NJ: Erlbaum.

Tinbergen, N. (1951). *The study of instinct.* New York: Oxford University Press.

Tooby, J. (1985). The emergence of evolutionary psychology. In Pines, D. (Ed), *Emerging syntheses in science,* pp. 106-122. Santa Fe: Santa Fe Institute.

Tooby, J. & Cosmides, L. (in press). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality.*

Tooby, J. & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In Kinzey, W. G. (Ed.), *The evolution of human behavior: Primate models,* pp. 183-237. New York: SUNY Press.

Wallace, A. R. (1904). *Man's place in the universe. (fourth edition).* London: Chapman and Hall.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought.* Princeton: Princeton University Press.

Williams, G. C. & Williams, D. (1957). Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution, 11,* 32-39.

Wilson, E. O. (1971). *The insect societies.* Cambridge, MA: Harvard University Press.

Wilson, E. O. (1975). *Sociobiology: The new synthesis.* Cambridge, MA: Harvard University Press.

ON WHY THERE ARE SO FEW COMPARISONS IN COMPARATIVE PSYCHOLOGY

Brian Mackenzie
University of Tasmania

ABSTRACT: The comparative study of behavior requires close attention to the ecologically unique details of the environmental challenges and adaptations (both behavioral and structural) of a systematically selected range of species. It offers an understanding of which aspects of behavior change and which remain constant across phylogenetic pathways and evolutionary challenges. The General Process View of Learning (also known as the principle of the transsituationality of reinforcement, and by several other names), however, militates against study of the details of behavioral adaptations, by insisting that particular behaviors may be regarded as arbitrary instances of universal associative principles. The history of behaviorism, and of contingency theory, in particular, is largely the history of the gradual emergence and dominance of this General Process View, and of the working out of its profoundly anticomparative implications. The increasingly wide repudiation of the General Process View is providing the basis for a renewal of comparative studies of behavior.

Comparative psychology, like psychology as a whole, has always attracted its share of commentators who are willing to pass judgment on its past record, its present health, and its future prospects. At the present time, and for quite a number of years in the past, the commentators have been split between those who find comparative psychology in robust health and those who find it in decline or worse.

Those who find comparative psychology healthy point to the amount of research in progress on the behavior of nonhuman species, to new or greatly amended behavioral principles that have come from this research during the last two decades, to the build-up of information about the adaptations of particular species, or to the use of animal data in developing general behavior theories (e.g., Dewsbury, 1984; Galef, 1987). The orientation of these positive commentators tends to be more behavioral than cognitive, more laboratory-based than field-based, and more operant than ethological. They tend to view contingency theory on a continuum ranging from tolerance to enthusiastic approval.

On the other hand, those who find comparative psychology less healthy, or altogether moribund, point to the limited number of different

---

Address correspondence to Brian Mackenzie, Department of Psychology, University of Tasmania, GPO Box 252C, Hobart, Tasmania 7001, Australia.

189

species typically studied in psychological laboratories (Porter, Johnson, & Granger, 1981), to the limited range of environments typically studied in animal experimentation (Doré & Kirouac, 1983), to the paucity of interspecies comparisons (Tolman, 1987), and to the absence of any meaningful, overarching theoretical goal of the discipline as a whole (Adler, Adler & Tobach, 1973; Lockard, 1971). These commentators are not united in their orientations, but many of them tend to be more cognitive than behavioral, more field-based than laboratory-based, and more ethological than operant. They tend to view contingency theory on a continuum ranging from tolerance to severe disapproval.

The two groups of commentators sometimes seem to be passing judgment on two quite different disciplines, and in an important sense I suspect they are. They have, I suggest, quite different conceptions of what comparative psychology should be, and thus have different criteria for assessing the current scene. At the descriptive level both groups are accurate; modern comparative psychology is characterized by both microtheoretical ferment (e.g., in the operant anlaysis of "foraging") and macrotheoretical stagnation (e.g., in the relative inattention paid to phylogenetic integration). Judgments about the health of the discipline depend on which of these is considered more important. I will not try to survey both groups of conceptions. Instead, I will present one conception which owes something to both groups, although it probably has more affinity with the "pessimistic" group. It is a relatively bald statement of what comparative psychology's mission should be, and how it started off fulfilling that mission, and how it lost sight of it, and why that was regrettable, and how it may have started to find its way again in recent years. The view is my own, but it has some significant points in common with that of Adler *et al.* (1973), Tolman (1987), and others. It also has some major divergences from those views, especially in its conclusions.

Comparative psychology, in this view, is not merely the study of the behavior of nonhuman organisms, or even the study of a specified subset of that behavior as Dewsbury (1984) has argued. It is instead the study of the evolutionary development and progression of behavior in all three of its major aspects: cognitive, affective, and conative. The subject matter of the discipline is neither rats nor pigeons nor human beings, but rather behavioral systems and the way they both change and remain constant across phylogenetic pathways and evolutionary challenges. A single species or group of species no more provides the subject matter for the field than it does for comparative anatomy, although a group of related species may comprise a specialist area in each. The understanding of human behavior is a legitimate goal (although not the only goal) of comparative psychology, just as the understanding of human anatomy is a legitimate goal of comparative anatomy. But progress toward achieving that goal in each case requires seeing the human subjects as comprising a node in a phyletic network, and cannot be achieved by regarding

them purely on their own. Comparative psychology, like comparative anatomy, may well therefore be "anthropotelic," taking the understanding of humans as its goal or end point, but can never be anthropocentric.

That is how I conceive comparative psychology, and I believe that that is how it was conceived by the pioneers in the field, such as Romanes, Hobhouse, Morgan, and Thorndike, working from the 1880s to the period just before the First World War. They all studied animal behavior, especially animal learning, not just to learn about the behavior or mind of the species they were investigating, and certainly not to learn about learning as a general process, unrelated to any particular species or environments, but explicitly to construct a behavioral phylogeny, so that human thought and action could be seen as further examples and developments of the patterns of adaptation they were discovering in their animal subjects. The goal of their studies was most often to understand human thought and action, so they were in this sense anthropotelic. In the beginning, especially in the work of Romanes, the descriptions of animal thought and action were based on those of humans, so they were clearly anthropomorphic, and this was a flaw that had to be overcome by Romanes' successors. But these researchers never lost sight of their goal of placing human thought and action firmly in an evolutionary context, and so they were never anthropocentric.

The collapse of this program was gradual over the first 10-20 years of the development of behaviorism after 1913. There is a nice irony in that behaviorism, which first developed as a response to conceptual problems in comparative psychology (cf. Mackenzie, 1977, Ch. 3), soon abandoned the comparative orientation of its parent discipline.

There were two main moves in this behaviorist repudiation of its own comparative background. The first was the rejection of instinct. Instinct had been a major organizing concept for comparative psychology from the beginning. However, the excessive use of the instinct doctrine in the 1920s (Tolman, 1923) encouraged behaviorist theorists to deny them as unobservable and gratuitous concepts, much as had been done with the concepts of mind and consciousness earlier. In the absence of a more sophisticated concept for describing species-typical behavioral adaptations, however, the loss of the instinct concept made many comparative questions difficult even to formulate.

The more important influence, however, was the development in behaviorist psychology of what Seligman (1970) has called the General Process View of Learning. In this view, the process of learning can be studied as easily and effectively in one species as in another, and as easily with one problem situation, response, reward (or later, reinforcer), and environment as another. As Seligman put it,

In instrumental learning, the choice of response and reinforcer is a matter of relative indifference; that is, any emitted response and any reinforcer

can be associated with approximately equal facility, and a set of general laws exist which describe acquisition, extinction, discriminative control, generalization, etc., for all responses and reinforcers. (Seligman, 1970, p. 407).

This principle, which Seligman also called the assumption of equivalence of associability, was earlier identified by Paul Meehl (1950), who called it the principle of the transsituationality of reinforcement. Both Seligman and Meehl concentrated on the generality of the general process view across situations, although Seligman's illustrations underlined its generality across types of organisms also. I also gave it a name somewhat later, in the specific context of Skinner's system, explicitly dividing the principle into what I called the assumptions of environmental and speciational generality (Mackenzie, 1977). More importantly, the principle had its origins in British associationist theory from the time of Hobbes (as discussed by Deese, 1965), in the notion that any idea could, by association, lead to any other idea. The principle, in one form or another, had thus been central to associative theories of learning and cognition since long before the introduction of evolutionary theory. It is also quite clearly incompatible with evolutionary theory, or with any theory which emphasizes unique patterns of adaptation for different groups of organisms. Consistent with this judgment, it was not emphasized by the early comparative theorists I mentioned, except eventually for Thorndike. Thorndike (1911) revived it from the later British associationist writers and proposed it as a major theoretical principle that any behavior could be connected with any "satisfier," as he called reinforcers. From Thorndike the principle passed into currency in later learning theory, but more as an assumption than as a theory.

In whatever form, the principle is profoundly anticomparative. It militates against any detailed comparative analysis of the problem situation and adaptive behavior of different classes of organisms, by stipulating that they must be fundamentally the same, the result of associative or S-R bonds formed according to unvarying universal laws. Inevitably, therefore, it is the laws that will be of interest rather than the specific adaptations, or attempts to link the adaptations (in any way other than as instantiations of the general laws) across species. Acceptance of the principle as an assumption therefore made comparative analysis of behavior unnecessary (or irrelevant) for understanding behavior, just as the rejection of the instinct doctrine made such analysis difficult and methodologically suspect. If any behavior will do, it makes sense to concentrate on the most convenient ones. The result was the retreat from comparative studies, and the concentration of experimental effort on a restricted behavioral repertoire in a small number of species (preeminently rats and pigeons) that has been widely documented and decried from Beach (1950) to Grosset and Poling (1982).

Acceptance of the principle also made it quite obscure as to why anyone would want to study animal learning. The two major justifications advanced during the 1940s and 1950s for the detailed study of learning in rats and pigeons were: (a) to study the process of learning for its own sake, and (b) to determine laws of learning which could subsequently be applied to the behavior of humans. The first amounts to the study of behavior without any behaving organisms, and underlines the antievolutionary (and often antibiological) character of later behaviorist learning theory. The second restricts the study of animal behavior to those features of it which may be helpful in explaining human behavior, and thereby incorporates precisely that anthropocentrism which was absent from the earlier comparative psychology. That fault, at least, was a minor one however, because the application of the "laws of learning" to human behavior was always more of a hope for the future than a focus for the present.

Fortunately, the assumption of the equivalence of associability, and with it the whole of the general process view of learning, is not only wrong but widely seen to be wrong. It has been shown to be wrong both in learning theory in general (Seligman, 1970) and in the system where it was most highly developed, that of Skinner's operant psychology (Mackenzie 1977, 1984). It is now fairly generally accepted that responses and reinforcers cannot be connected indifferently, that in each species of organism some behaviors can be potentiated best by certain reinforcers and other behaviors by others; so that in the pigeon, key-pecking can be reinforced very well by food, and wing extension can be reinforced very well by shock avoidance, but neither response can be reinforced nearly as well by the other reinforcer. There seems almost to be some intrinsic connection between pecking and food, and between extending the wings and avoidance of danger, and of course there is. The breakdown of the principle of the equivalence of associability, in other words, is not adventitious; it does not just happen. Instead, it happens in a consistent and meaningful way, and to understand the way it happens it is necessary to understand the details of the organism's behavioral adaptation to its environment.

In short, the principle that was central to the behaviorist dismissal of the comparative study of behavioral adaptations can be seen to have failed, precisely because those adaptations impinge at every stage even upon the restricted behavior studied in the behaviorist laboratory. They impinge by placing constraints on that behavior, constraints that prevent the behavior from consistently following general laws of learning and that can be taken account of within the learning theory only by introducing them as *ad hoc* supplementary principles ("instinctive drift," according to Breland & Breland, 1961; a "dimension of preparedness," according to Seligman, 1970). Such supplementary principles, however,

rather than strengthening the learning theory accounts, merely lessen their applicability; they acknowledge the need for an analysis of behavioral adaptations in the natural environment as the basis for understanding behavior in the operant chamber, but do not provide it. Meehl (1950) was correct in rewriting the empirical law of effect to read "All reinforcers are transsituational"; and the failure of the assumption of the equivalence of associability amounts, quite simply, to the refutation of the law of effect.

Finally, I come to my optimistic conclusion. The refutation of the general process view of learning at the hands of Seligman and other writers has helped spark a revival of interest in comparative studies of animal adaptation and even animal minds. Animal cognition has emerged as a thriving field in the past 10-15 years, with the application to animal studies of some of the models and insights of the human cognitive laboratory (Roitblat, 1982; Roitblat, Bever, & Terrace, 1984). Other authors are again explicitly trying to understand behavioral adaptations in the context of broad descriptive evolutionary trends (Gottlieb, 1984, 1987). The important point is that it is specifically as part of the collapse of some of the central assumptions of associative learning theories that these signs of a rebirth of comparative psychology may be seen. If the influence of these theories continues to wane, and if the general process view of learning which underlay them gradually becomes recognized as intellectually bankrupt, then we may look forward once again to seeing widespread comparisons in comparative psychology.

REFERENCES

Adler, H. E., Adler, L. L., & Tobach, E. (1973). Past, present and future of comparative psychology. *Annals of the New York Academy of Sciences, 223*, 184-192.

Beach, F. A. (1950). The snark was a boojum. *American Psychologist, 5*, 115-124.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist, 16*, 681-684.

Deese, J. A. (1965). *The structure of associations in language and thought.* Baltimore: Johns Hopkins University Press.

Dewsbury, D. A. (1984). *Comparative psychology in the twentieth century.* Stroudsbury, Pennsylvania: Hutchinson Ross.

Doré, F. Y., & Kirouac, G. (1987). What comparative psychology is about: Back to the future. *Journal of Comparative Psychology, 101*, 242-248.

Galef, B. G. (1987). Comparative psychology is dead: Long live comparative psychology. *Journal of Comparative Psychology, 101*, 259-261.

Gottlieb, G. (1984). Evolutionary trends and evolutionary origins: Relevance to theory in comparative psychology. *Psychological Review, 91*, 448-456.

Gottlieb, G. (1987). The developmental basis of evolutionary change. *Journal of Comparative Psychology, 101*, 262-271.

Grosset, D., & Poling, A. (1982). The status of *Rattus:* Rats as subjects in articles published in *Journal of Comparative and Physiological Psychology* from 1969 to 1981. *Psychological Reports, 51*, 969-970.

Lockard, R. L. (1971). Reflections on the fall of comparative psychology: Is there a message for all of us? *American Psychologist, 26*, 168-179.

Mackenzie, B. (1977). Behaviorism and the limits of scientific method. New York: Humanities Press.

Mackenzie, B. (1984). The challenge to Skinner's theory of behavior. *Behavioral and Brain Sciences, 7,* 526-527.

Meehl, P. E. (1950). On the circularity of the law of effect. *Psychological Bulletin, 47,* 52-75.

Porter, J. H., Johnson, S. B., & Granger, R. G. (1981). The snark is still a boojum. *Comparative Psychology Newsletter, 1,* 1-13.

Roitblat, H. L. (1982). The meaning of representation in animal memory. *Behavioral and Brain Sciences, 5,* 353-406.

Roitblat, H. L., Bever, T. G., & Terrace, H. S. (Eds.). (1984). *Animal cognition.* Hillsdale, NJ: Erlbaum.

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review, 77,* 406-418.

Thorndike, E. L. (1911). *Animal intelligence.* New York: Macmillan.

Tolman, C. W. (1987). Comparative psychology: Is there any other kind? *Journal of Comparative Psychology, 101,* 287-291.

Tolman, E. C. (1923). The nature of instinct. *Psychological Bulletin, 20,* 200-218.

THE GENERAL PSYCHOLOGICAL CRISIS AND ITS COMPARATIVE PSYCHOLOGICAL RESOLUTION

Charles W. Tolman
University of Victoria

ABSTRACT: The crisis in general psychology is identified as one of theoretical indeterminacy. An important source of indeterminacy is the form of generalization that emphasizes classification and common characters and identifies the general with the abstract. Determinate theory requires a form of generalization that identifies the general with the concrete and emphasizes genesis and interconnection. In order to overcome its crisis, general psychology thus requires the kind of evolutionary methodology that comparative psychology is, historically speaking, best prepared to provide.

> Science must begin with that with which real history began. Logical development of theoretical definitions must therefore express the concrete historical process of the emergence and development of the object. Logical deduction is nothing but a theoretical expression of a real historical development of the concreteness under study. (Ilyenkov, 1982, p. 200)

The decade of the 1970s was a difficult period for comparative psychology. The troubles first became apparent somewhere around 1950 when Frank Beach called attention, in a deploring tone, to our "excessive concentration" on a very limited number of species, especially the ubiquitous *Rattus norvegicus.* It is now clear to me that these "troubles" did not in fact amount to a real "crisis." This, then, is as good a place as any to recant my earlier claim that they did so (Tolman, 1984).

In the same year, 1950, and for not dissimilar reasons, we were taken to task by Konrad Lorenz who found it "misleading if psychologists, who evidently are not familiar with what we (biologists) mean when we speak of the comparative method, apply the same term in a very loose sense to all behavior studies concerned with different forms of life" (p. 239). His reaction was a biting one: "I must confess that I strongly resent it, not only from the terminological viewpoint, but also in the interests of the very hard-working and honest craft of really comparative investigators, when an American journal masquerades under the title of 'comparative'

197

psychology, although, to the best of my knowledge, no really comparative paper ever has been published in it" (pp. 239–240). The doubts seemed to mount until 1969 when it was announced by Hodos and Campbell that the evolutionary aspirations of comparative psychology were misplaced and there could consequently be no comparative psychological theory. Lockard (1971) followed quickly to announce the "fall of comparative psychology." And then, whatever pieces might be left, Wilson announced in 1975, would soon be "cannibalized," along with our sister discipline ethology, by sociobiology.

The responses to all this were varied (e.g., Tobach, Adler, & Adler, 1973; Wyers *et al.*, 1980), but I should hope that by now most thoughtful observers will agree with Gottlieb (1984) that comparative psychology has survived, does have theory (if not *a* theory), and that this theory is potentially and genuinely evolutionary (cf. also Ardila, 1986, and Ribes, Ibañez, and Hernandez Pozo, 1986). The onslaught of the skeptics and disbelievers has done us the great favor of awakening us from our "dogmatic slumbers" (Kant). We are probably more keenly aware now of the evolutionary-theoretical potentials of comparative psychology than ever before in our history. I shall return to this line of discourse momentarily.

THE CRISIS IN GENERAL PSYCHOLOGY

In the meantime, while comparative psychology was experiencing its own peculiar difficulties, general psychology was going through what was recognized both inside and outside the discipline as a genuine "crisis." As one psychologist put it: "Whether they (psychologists) are experiencing an identity crisis, a paradigmatic crisis, or a crisis of confidence, most seem agreed that a crisis is at hand . . ." (Elms, 1975, p. 968). In his introduction to the 1975 Nebraska Symposium on Motivation, each of the eight contributions to which was in its own way crisis-oriented, the editor, William Arnold, expressed his understanding of the matter as follows:

> To the extent . . . that psychology has been committed to an older, Newtonian conceptual system (and its interlocking philosophical assumptions), it became apparent that we, in our domain (psychology), have been guiding ourselves by a limited conception of science and, accordingly, by a restricted conception of the human being. Nonetheless, in recent years many psychologists have been acquiring a sharper perception of the mutable, evolutionary character of all theoretical conceptions—including the regnant conceptions in psychology. As a consequence psychology appears to be in a period of transition or "crisis." (Arnold, 1976, p. vii).

Two implications of this statement should be examined a little more closely before going on to specify the exact nature of this crisis. First, Newton is getting here a little more of the blame than he deserves. The problem was not the "Newtonian conceptual system" as such, but rather the 19th-century positivist version of that system. The importance of this distinction will become more evident as we see the specific nature of the crisis more clearly. Second, it is worth noting that the alleged insight regarding the "evolutionary" character of psychological concepts is not elaborated by Arnold. It turns out, however, by the present analysis, to be highly prophetic.

So, what is the precise nature of the crisis? I will present here quotations from several participants in the crisis debates in order to show that a consensus on the focus does in fact emerge. First, Amedeo Giorgi, a contributor to the 1975 Nebraska Symposium:

> Clearly, there is a lack of unity in psychology with respect to theories, methods, importance of data, definition of subject matter and almost every other important dimension. Moreover, the same critiques concerning the science of psychology perseverate. Why? Because somehow the organization of the field of psychology is not meeting the pervasive needs of all of its practitioners. And this, in turn, is because its organization has not yet found the central viewpoint for integrating the various facets of its field. So one meaning of lack of unity in psychology is the fact that the point of view or central perspective for integrating the disparate aspects of psychology has not yet been achieved or elaborated. This means that the perspective adopted is exhausted before one comes to the end of commonly accepted psychological phenomena. Or it means that one can comprehend the totality of psychological phenomena only by excessive abstraction or ambiguous labeling (Giorgi, 1976, p. 285).

Although his main concern is with social psychology, the assessment of the crisis by Moscovici (1972) is similar:

> The respect of common sense, the proliferation of experimental studies lacking theoretical preoccupations, and the isolation of various areas of research in social psychology combine to explain the accumulation of facts and notions which do not amount to real progress since they are not conceptually integrated and since no theory is, in any real sense, disconfirmed or replaced by another. The concepts employed have their origin in other fields; theoretical models exist side by side in a relationship which neither constitutes real dialogue nor fertile contradiction. It is therefore not surprising that the empirically established facts are nothing but a heterogeneous collection, as are the theories on which they are supposed to depend. The experiments and empirical studies are not really capable of confrontation in a common framework . . . . (Moscovici, 1972, pp. 43-44)

Catania (1973) was obviously seeing the same state of affairs when he wrote:

Students of psychology still are asked to choose theoretical sides. They see functional accounts of operant behavior pitted against ethological accounts of behavioral structure, analyses of reinforcement contingencies pitted against theories of cognitive processing, and descriptions of language as verbal behavior pitted against psycholinguistic formulations of language competence . . . psychologists are not yet even agreed on whether theirs is a science of behavior or science of mental life. (Catania, 1973, p. 434)

The focus that emerges in these and many other assessments is succinctly captured in the title of Arthur Staats' 1983 book, *The Crisis of Disunity in Psychology*. It appears that psychologists suddenly found themselves without any real basis for deciding among theoretical alternatives. Even the traditional basis in "facts" or data appeared to have abandoned them. This was recognized already in 1966 by Hilgard and Bower who observed then that: "One of the most perplexing problems within psychological research is how to make the gains cumulative, that is to build a firm foundation from past research on which to plan towards the future" (1966, p. 582). After rehearsing a litany of problems, such as latent learning and peripheral versus central mediation, that had not been resolved by "crucial" experiments, they concluded that: "Accumulation of knowledge means neither mere fact-gathering nor isolated hypothesis-testing, but thoughtful systematic approaches to meaningful questions leading to *conclusive thinking*" (1966, p. 583).

This need for an adequate theoretical solution to the crisis was echoed in 1972 by Harré and Secord, whose remarks, though directed at social psychology, apply just as well to general psychology:

The need for a comprehensive theoretical treatment of social psychology and for a reformed methodology we feel to be pressing, and to be evident from the increasing dissatisfaction with the state of social psychology, even within the citadels of the profession. The underlying reason for this state we believe to be a continued adherence to a positivist methodology, long after the theoretical justification for it, in naive behaviourism, has been repudiated. At present there is scarcely any coherent body of theory. In such a vacuum it is still possible to carry on empirical studies which make sense only if people are conceived of in the mechanical tradition as passive entities whose behaviour is the product of 'impressed forces,' and whose own contribution to social action is the latent product of earlier impressed experience. A methodology of experiment survives in which the typical investigation is recommended to be a manipulation of 'variables,' and the typical result a correlation in the manner of Boyle's Law. (Harre & Secord, 1972, p. 1)

There is remarkable agreement among the assessments sampled that the general nature of the crisis is one of lack of unity. I understand this to mean a lack of coherence, a lack of agreement among psychologists about what counts. The principal symptom has been designated as "theoretical

indeterminacy" or as the "Beliebigkeit psychologischer Theorien," as it has become known in a particular segment of the German psychological literature (Holzkamp, 1978). We are faced with numerous theories of personality, motivation, learning, social influence, development, etc., with no apparent rational grounds for choosing from among the alternatives or resolving their differences, and therefore with no basis for the progressive accumulation of coherent psychological knowledge.

It should not be imagined here that all psychologists have been equally distressed by this situation. A small, but influential number of psychologists has taken it as the necessary state of affairs, and has even welcomed it openly as a relief from what Sigmund Koch called "single-principle imperialism" (Koch, 1981). Other examples of such advocacy of pluralism are found in the work of Dixon (1983), Gergen (1981), and Royce (1982). Even Hilgard appears to have retreated toward some kind of pluralism (Hilgard, 1987, p. 803). The major difficulty with all these versions of metaphysical pluralism is that they lead logically, and therefore necessarily, to nihilism and solipsism. As I shall attempt to show in a moment, we need not accept this state of affairs as necessary; there *is* a solution to the problem of theoretical indeterminacy.

A second important conclusion that can be drawn from the crisis debates is that the solution is not a strictly empirical one. Perhaps the most important lesson to be learned from this historical development is that the accumulation of coherent knowledge is not achieved through mere induction from facts or data. The "crucial experiment" (i.e. an appeal to more data—presumably of the "right" kind—for an automatic resolution to theoretical differences) is a myth. While the right data are indeed important, even indispensable, they are not sufficient in themselves to resolve theoretical indeterminacy, thereby overcoming the crisis. As Hilgard and Bower put it in 1966, we need "conclusive thinking," whatever that is. This is the point to which we shall return in a moment.

A third conclusion is one that is essential to our understanding, though it will not preoccupy us here. This is that the philosophical root of the problem is not Newtonian materialism, but the so-called "logical" form of 19th-century positivism. It is well known that one of the main goals of modern positivism was the unification of science under some all-embracing theoretical framework. It is equally well known that all attempts at this ended in miserable failure. The reason was positivism's inability to resolve, within the framework of its own assumptions, the problem of verification (Passmore, 1967). Now the problem of verification is precisely the problem of theoretical determinacy. At the bottom of this difficulty with verification is the separation of the thing-as-known from the thing-in-itself inherited from Humean empiricism via Kantian idealism. The lesson from all of this is that an adequate metaphysical foundation for scientific thinking is, at the very least, one that grants the

access of cognition to objects themselves. The only way in which we can comprehend our agreement that the grass is green is by appeal to the fact that the grass *is* green, with the latter state of affairs having priority over the former. But it is now clear from our second conclusion that this kind of "direct realist" empiricism, while essential, will not alone provide the solution we are seeking. What, then, is the additional theoretical/methodological requirement for an acceptable solution?

TOWARD A RESOLUTION OF THE CRISIS

It is important to note at the outset that the solution I will describe is not new. It is an understanding of scientific activity that is evident in the behavior of scientists beginning at least with Galileo. The fact that crises like psychology's have not occurred in physics or biology is, I believe, largely a reflection of this. It is an understanding however, that has been neither recognized—at least as such—nor articulately developed in English-language philosophy of science. In short, there has come about a discrepancy between the practice and theory of science that has effected more mischief in psychology than in any other science (although deleterious effects are noticeable to one degree or another in all of the social sciences, for which there are good historical reasons, which we are unable to develop here).

The problem, and its solution, lies in our understanding of the process of generalization that forms the basis of all theory construction. What our theory of scientific practice has failed to recognize is that there are two distinct, broad forms of generalization: one which emphasizes classification and common characters, and identifies the general with the abstract; and a second which emphasizes the discovery of genesis and interconnection, and identifies the general as concrete (actually "concrete universal"). It is the first that has almost exclusively dominated theory-building efforts in psychology since the early part of this century. This empirical generalization to the abstract seeks to identify traits (or principles) that are common to the largest number of instances, aiming of course at universality. Davydov (1984, pp. 20-21) cites six characteristics of "empirical knowledge," which is his term for the product of this form of generalization:

1. Empirical knowledge is produced by comparing objects and their representations which makes it possible to discern in them the same general traits.

2. Comparison discerns the formally common trait which makes it possible to classify separate objects under a certain formal class irrespective of their being related to each other.

3. Empirical knowledge, which is based on observation, reflects only external traits of objects and for this reason completely relies on perceptual conceptions.

4. The formally common trait is equal to the individual traits of objects.

5. The concretization of empirical knowledge consists in gathering of illustrations or examples which belong to a formally derived category.

6. The necessary means of crystallizing empirical knowledge is the word or the term.

We will not take the time to work through examples here. It should be fairly evident that these characteristics apply to such categories as "reinforcement." Just consider all the concrete detail that must be ignored in order to achieve the claimed universality of this notion. Note also that this category refers not only to a trait or character of learning, but has been advanced as a *theory* of learning. (It is instructive to note here that while the "law of reinforcement" is an empirical generalization, the "law of falling bodies" is not.) The psychology of personality and motivation are replete with examples of this sort. Whole theories of personality have been advanced on the basis of "self-esteem," while whole theories of motivation have been advanced on the basis of "achievement" or "social comparison."

There is no denying that this kind of abstraction and generalization has its place in scientific work. The difficulties arise when it is taken as the principal or only form of generalization. When this occurs, it becomes the main source of theoretical indeterminacy, which, it will be recalled, I have identified as the cause of the crisis in general psychology. How does this happen? There are at least two ways. First, there are no constraints within the abstraction process by which to decide what is to be abstracted. Any class of objects is likely to yield any number of equally "good" abstractions. This is undoubtedly the case in personality theory with its endless lists of traits. What's more, the formal criteria can lead to results that are clearly misleading. For example, we might abstract from the class "mankind" both the trait of soft earlobes and toolmaking. In fact the former is far more general, indeed it is universal, yet hardly anyone with a modern scientific understanding would be willing to conclude that soft earlobes are more essential (because universal) to being human than toolmaking. At any rate, it is clear that any number of competing theories can exist on this basis about just about anything. They may all be equally "true," and there appears to be no criterion by which to select among them. One seeming solution is the eclectic one, but this cannot be a real solution as it does not overcome the basic difficulty.

The second source of indeterminacy in this form of generalization is a circularity that, given no additional guides, can only be resolved arbitrarily. This method of generalization cannot actually discover the characters of any particular category because it must first define the category in order to decide what is general to it. The abstraction and generalization process cannot, in short, define the category. If, therefore, initial definitions cannot be determined by the scientific investigative process (generalization), then they are necessarily arbitrary, and noth-

ing prevents different investigators from beginning with different definitions. Then, owing to the different definitions, disparate lists of general traits will perforce be found. Again, there is nothing in the process of generalization to which one can appeal in order to find a rational resolution of these differences.

The second form of generalization, called "substantial generalization" by Davydov (1984), overcomes these difficulties by allowing the object or category to define itself. That is, it contains within it the means by which initial definitions can be corrected in the process of investigation. According to Davydov:

> Substantial generalization reveals the essence of objects in the form of developmental laws showing what defines their development. The object of this type of generalization is to reveal a law which is the necessary connection of specific phenomena within a whole, the law of the origin of this whole. The revelation of the general nature of a real relation takes place in the analysis of such specific features which allow it to be a genetic basis of a developed system. The beginning of the concretization of these particular phenomena is the essence of the discovery of the universality of the relation in question (Davydov, 1984, pp. 12-13).

The word "genetic" is used here in its original sense as "pertaining to the origin, history, and development of an organism" (English & English, 1958, p. 223).

But by what criterion do we recognize that we have achieved such generalization? The answer is straightforward: it is the answer to the question "whether the particular phenomenon directly expressed in it is at the same time the universal genetic basis from the development of which all other, just as particular, phenomena of the given concrete system may be understood in their necessity" (Ilyenkov, 1982, p. 76). It is thus only through this form of generalization that we can claim that toolmaking is more essential to human nature than are soft earlobes. "Toolmaking" is an induction that guarantees deduction of many specifically human traits, from the form of the human jaw to a complex division of labor (see Woolfson, 1982). "Toolmaking" helps us to understand all the important ways in which humans are different from other animals. Soft earlobes certainly do not.

But this points up a very special feature of the concrete universal which is the result of substantial generalization. This is that it need not be empirically, statistically universal. There are humans who do not make tools. What makes toolmaking universal for the human species is that even what is done by those who do not make tools is understandable in terms of our toolmaking origins. It is its function as a "genetic basis" that makes toolmaking universal. And it is precisely this that makes the abstraction "toolmaking" concrete: it aids in revealing the "necessary connection of specific phenomena within a whole" (Davydov, 1984, p. 13).

Davydov offers the following characteristics of the "theoretical knowledge" that results from substantial generalization to the concrete universal, as contrasted with merely empirical knowledge:

1. Theoretical knowledge arises on the basis of an analysis of the role and function of a certain relation of things inside a structure system.

2. By means of the analysis, this real, particular relation of things is searched for which at the same time is the genetic basis of all other manifestations of the system. This relation appears as a universal form or the essence of mentally reproduced totality.

3. Theoretical knowledge which is based on the transformation of objects reflects their internal relations and interconnections. In the reproduction of an object, in the form of theoretical knowledge, thinking exceeds the limits of perceptual presentations.

4. In theoretical knowledge, the connection between the real universal relation and its various manifestations, i.e., the connection of the general and the individual, is crystallized.

5. The concretization of theoretical knowledge presupposes its conversion into a developed theory by the deduction and explanation of the specific manifestations of the systems relying on its universal basis.

6. Theoretical knowledge appears primarily as methods of intellectual activity and subsequently in different symbolic sign systems in which for example artificial and natural language is used (a theoretical concept may appear as a method of deducing the specific from the universal, but still lack a corresponding terminological formulation). (Davydov, 1984, pp. 20-21)

Research on evolutionary origins is the prime example of substantial generalization. Making this identification allows us to see that evolutionary research is not just another form of research, it is the model for all good research. It is only by revealing the evolutionary or other genetic origins of an object that a determinate theory about it can be constructed. Only in this way can we decide whether earlobes or hand structure are essential to our theory of human nature.

IMPLICATIONS FOR COMPARATIVE PSYCHOLOGY

The general implications for psychology are obvious. The crisis of indeterminacy exists because of a one-sided reliance on empirical generalization to the abstract. The solution is therefore to develop a methodology based upon substantial generalization to the concrete. This methodology, at least in its broad strokes, already exists in evolutionary and other genuinely genetic research. The psychological subdiscipline most prepared to carry out such a project is comparative psychology. In order for comparative psychology to meet the objective which I am proposing, two broad reforms will be necessary: first, it will have to adopt a stronger commitment than it has shown in recent history to a truly evolutionary, genetic methodology; and, second, it will have to expand its

conception of its mission to include a central commitment to providing the foundation for a determinate (and in this sense unified) general psychological theory. This is, as I have pointed out elsewhere (Tolman 1987a, b), much more in line with the original intentions of our discipline's founders, namely Romanes, Morgan, and Hobhouse, than with the arid psychologies of Thorndike and Watson.

I see the prospects for all of this as quite auspicious. There are at least three grounds, besides the theoretical one discussed above, for my optimism. The first is the "state of the art" of comparative psychology. As I have already indicated, we have survived our troubled times and have even come out of them a good deal wiser. We have been forced to take a close look at our theories and methods, to become more explicit about them, and to pay closer attention to their evolutionary nature. The discussions on anagenesis and levels (e.g., Aronson, 1984; Yarczower, 1984) have been, in my view, an important part of our reexamination; and the post-neoDarwinian thinking on evolutionary process may prove to be even more significant (Ho, 1987), especially when combined with a recognition of the leading role of behavior in evolution (e.g., Napier, 1976, p. 4). It is conceptions like these that are enabling us better to formulate our evolutionary, genetic, and therefore "substantial," mission.

An emphasis on continuity in evolution served a very important historical purpose in the late 19th and early 20th centuries, but it goes awry when we allow it to back us into the reductionist corner from which it appears that animal species are not really different and that nothing psychological has actually evolved. The concepts of anagenesis and levels help us to restore discontinuity, now scientifically conceived, to its rightful place. It is, after all, discontinuity that evolutionary theory must explain, not continuity. Armed with such methodological notions, comparative psychology is better than ever prepared to face up to the challenge of explaining why humans are different from apes, why apes are different from rats, why vertebrates are different from invertebrates, and so on, without fear of falling into the error of the *Scala naturae*, rather than spending all its energy on abstract notions that supposedly unite all the species, yet in fact explain nothing.

A second ground for my optimism is that we are primates. According to Napier:

> Primates are unique among the living mammals in displaying the 'staircase' phenomenon. In other orders the evolutionary sequence shows the successive replacement of archaic forms by new and improved ones. Take for example the horse. There is only one genus of horse today; there is no *Eohippus, Merychippus,* or *Pliohippus* to provide the milestones on the road to *Equus.* As each 'new' horse evolved, the 'old' horses went out of business; not so with the primates. As each new grade of primates evolved, the old grades, which had established themselves in their own particular environmental niches, continued to flourish." (Napier, 1972, p. 6).

In other words, the anagenetic grades of most immediate concern for human evolution are represented by living species. Add to this the rapidly expanding information now available from paleoanthropology, especially about *Australopithicus*, and one could hardly imagine a more advantageous position from which to develop a comparative psychology of the human species. I hasten to add, however, that this in no way disparages the important work on preprimate grades of psychological evolution. Primates, as an order, are quite advanced and that, too, must be explained.

The third ground for my optimism is the fact that a comprehensive scheme of psychological evolution has already been developed, which, it appears to me, holds much promise as a guiding thread for an evolutionarily reconstructed comparative psychology. This is the scheme developed in the 1940s and 50s by A. N. Leontyev (1981) and his colleagues at Moscow University as a foundation for current "activity theory" and recently developed further by Holzkamp (1973) and Schurig (1975a, 1975b, and 1976). In anagenetic terms, the scheme outlines the major grades of what Leontyev chose to call "psychic" development. It is a scheme that has the enormous advantage of being thoroughly ecological in outlook (i.e. it does not deal with the evolution of merely psychic qualities independent of biological qualities or of the environment, but is guided by a clear idea of essential subject–object interrelations). It also contains hypotheses about the causes of intergrade transitions. In short, it is a scheme that is totally in keeping with the demands of substantial generalization. As I have outlined the details of the scheme elsewhere (Tolman, 1987c), I shall not say more about it here.

CONCLUSION

I have shown that the crisis in general psychology is such that it can only be resolved by resorting to an alternate form of generalization that leads to concrete theoretical knowledge rather than abstract generalities. This alternative form of generalization turns out to be one that is evolutionary (genetic or developmental) in nature. In a broad sense, general psychology, if it is to overcome its difficulties, must become a truly evolutionary (genetic or developmental) psychology. Although comparative psychology has had historical growing pains of its own, we now see it moving toward the development of precisely the kind of methodology that is lacking in general psychology. It is in this sense, then, that the *resolution* of the general psychological crisis is *comparative* psychological. Its evolutionary methodology is the only possible basis for a theoretically healthy general psychology. One can easily imagine that as general psychology becomes more and more informed by comparative psychological methodology, the distinction between the two psycholo-

gies will diminish: the discipline will rightly become more like its sister, biology, in which comparative biology is the only kind that exists.

REFERENCES

Ardila, R. (1986). Significado y necesidad de la psicologia comparada. *Revista Latinoamericana de la Psychologia, 18,* 157-169.

Arnold, W. J. (1976). Introduction. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation 1975.* Lincoln, NE: University of Nebraska.

Aronson, L. R. (1984). Levels of integration and organization: a reevaluation of the evolutionary scale. In G. Greenberg & E. Tobach (eds.), *Behavioral evolution and integrative levels,* (pp. 57-81). Hillsdale, NJ: Erlbaum.

Beach, F. A. (1950). The snark was a boojum. *American Psychologist, 5,* 115-124.

Catania, A. C. (1973). The psychologies of structure, function and development. *American Psychologist, 28,* 434-443.

Davydov, V. V. (1984). Substantial generalization and the dialectical-materialist theory of thinking. In M. Hedegaard, P. Hakkarainen, & Y. Engeström (Eds.), *Learning and teaching on a scientific basis.* Aarhus, DK: Aarhus Universitet.

Dixon, R. A. (1983). Theoretical proliferation in psychology: a plea for sustained disunity. *The Psychological Record, 33,* 337-340.

Elms, A. C. (1975). The crisis of confidence in social psychology. *American Psychologist, 30,* 967-976.

English, H. B., & English, A. C. (1958). *A comprehensive dictionary of psychological and psychoanalytic terms.* New York: Longmans, Green & Co.

Gergen, K. J. (1981). The meager voice of empiricist affirmation. *Personality and Social Psychology Bulletin, 7,* 333-337.

Giorgi, A. (1976). Phenomenology and the foundations of psychology. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation 1975,* (pp. 281-348). Lincoln, NE: University of Nebraska.

Gottlieb, G. (1984). Evolutionary Trends and evolutionary origins: relevance to theory in comparative psychology. *Psychological Review, 91,* 448-456.

Harré, R., & Secord, P. F. (1972). *The explanation of social behaviour.* Oxford: Basil Blackwell.

Hilgard, E. R. (1987). *Psychology in America: a historical survey.* New York: Harcourt Brace Janovich.

Hilgard, E. R., & Bower, G. H. (1966). *Theories of learning.* New York: Appleton-Century-Crofts.

Ho, M. W. (1987). Evolution by process, not by consequence: implications of the new molecular genetics on development and evolution. *International Journal of Comparative Psychology, 1,* 3-27.

Hodos, W., & Campbell, C. B. G. (1969). *Scala naturae:* why there is no theory in comparative psychology. *Psychological Review, 76,* 337-350.

Holzkamp, K. (1973). *Sinnliche Erkenntnis—Historischer Ursprung und gesellschaftliche Funktion der Wahrnehmung.* Frankfurt a. M.: Athenäum Verlag.

Holzkamp, K. (1978). Die Überwindung der wissenschaftlichen Beliebigkeit psychologischer Theorien durch die Kritische Psychologie. In Holzkamp, K., *Gesellschaftlichkeit des Individuum.* Köln: Pahl-Rugenstein.

Ilyenkov, E. V. (1982). *The dialectics of the abstract and the concrete in Marx's Capital.* Moscow: Progress.

Koch, S. (1981). The nature and limits of psychological knowledge. *American Psychologist, 36,* 257-269.

Leontyev, A. N. (1981). *The problems of the development of the mind.* Moscow: Progress.

Lockard, R. B. (1971). Reflections on the fall of comparative psychology: is there a message for us all? *American Psychologist, 26,* 168-179.

Lorenz, K. Z. (1950). The comparative method in studying innate behaviour patterns. In *Physiological mechanisms in animal behaviour,* (pp. 221-268). Cambridge: University Press.

Moscovici, S. (1972). Society and theory in social psychology. In J. Israel & H. Tajfel (Eds.), *The context of social psychology,* (pp. 17-68). New York: Academic Press.

Napier, J. R. (1972). *Primates and their adaptations.* Oxford: Oxford University Press.

Napier, J. R. (1976). *Primate Locomotion.* Oxford: Oxford University Press.

Passmore, J. (1967). Logical positivism. In P. Edwards (Ed.), *The encyclopedia of philosophy.* New York: Macmillan.

Ribes, E., Ibañez, C., & Hernandez Pozo, R. (1986). Hacia una psicologia comparativa: algunas consideraciones conceptuales y metodologicas. *Revista Latinoamericana de Psicologia, 18,* 263-276.

Royce, J. R. (1982). Philosophic issues, Division 24 and the future. *American Psychologist, 37,* 258-266.

Schurig, V. (1975a). *Naturgeschichte des Psychischen 1: Psychogenese und elementare Formen der Tierkommunikation.* Frankfurt a. M.: Campus Verlag.

Schurig, V. (1975b). *Naturgeschichte des Psychischen 2: Lernen und Abstraktionsleistungen bei Tieren.* Frankfurt a. M.: Campus Verlag.

Schurig, V. (1976). *Die Entstehung des Bewusstseins.* Frankfurt a. M: Campus Verlag.

Staats, A. W. (1983). *Psychology's crisis of disunity: philosophy and method for unified science.* New York: Praeger.

Tobach, E., Adler, E., & Adler, L. L. (1973). Comparative psychology at issue. *Annals of the New York Academy of Sciences, 223.*

Tolman, C. W. (1984). Krize srovnávaci psychologie, pojem úrovni a nutnost dialektiky. *Ceskoslovenská* Psychologies, 28, 1-6.

Tolman, C. W. (1987a). Comparative psychology: is there any other kind? *Journal of Comparative Psychology, 101,* 287-291.

Tolman, C. W. (1987b). Zur Vorgeschichte der historischen Heransgehensweise in der bürgerlichen Psychologie. In W. Maiers & M. Markard (Eds.), *Kritische Psychologie als Subjecktwissenschaft,* (pp. 228-240). Frankfurt: Campus Verlag.

Tolman, C. W. (1987c). The comparative psychology of A. N. Leontyev (U.S.S.R.). In E. Tobach (Ed.), *Historical perspectives and the international status of comparative psychology,* (pp. 203-209). Hillsdale, NJ: Erlbaum.

Wilson, E. O. (1975). *Sociobiology.* Cambridge MA: Harvard.

Woolfson, C. (1982). *The labour theory of culture.* London: Routledge & Kegan Paul.

Wyers, E. J., Adler, H. E., Carpen, K., Chizsar, D., Demerest, J., Flanagan, O. J., von Glasersfeld, E., Glickman, S. E., Mason, W. A., Menzel, E. W., & Tobach, E. (1980). The sociobiological challenge to psychology: on the proposal to "cannibalize" comparative psychology. *American Psychologist, 35,* 955-979.

Yarczower, M. (1984). Behavior and evolutionary progress: anagenesis, grades, and evolutionary scales. In G. Greenberg & E. Tobach (Eds.), *Behavioral evolution and integrative levels,* (pp. 105-119). Hillsdale, NJ: Erlbaum.

Instructions to Authors

All manuscripts are reviewed by two readers other than the Editors. Authors should send three copies of their manuscripts to the National Editor of their country. If there is no editor, the paper should be sent to the Editor (Tobach). See inside front cover for editors' addresses. The manuscript should have a cover page with the full title, the author(s) name(s) and address(es) and should indicate the person with whom the Journal may correspond. The abstract should be on a separate page. All articles are published in English; the author(s) may also submit an abstract in another language which does not require non-Roman type. Author(s) are responsible for the accuracy of the non-English abstract. Acknowledgements should appear on a separate page and carry the title of the article and the author(s) name(s). If the author(s) prefer(s) anonymity in the review process, the manuscript, abstracts and illustrative material should carry only the title of the article. The cover sheet and acknowledgements will not be sent to the reviewers in this instance.

GENERAL REQUIREMENTS: Logical organization is essential. While headings help to structure the content, titles and headings within the manuscript should be as short as possible. Tables and figures should not be redundant. Illustrative material should be self-explanatory and used spar ingly. Tables and figures must be in camera-ready condition and include captions. Only items cited in manuscripts should be listed as references. Page numbers must be provided for direct quotations. Footnotes should not be used except in special circumstances and with permission of the Editor.

MANUSCRIPTS should be double spaced, the original on white bond paper, with margins of 3cm on all four edges and approximately 26 lines to the page. Manuscripts longer than 20 typed pages (not counting figures, tables and references) will only be considered in unusual circumstances. The style manual of the American Psychological Association (3rd Edition, 1983) is to be followed. If the manual is not available, authors may write to the National Editor or the Journal Editors (Tobach or Thompson) for a style sheet.

REFERENCES should be listed on a separate page and referred to in the text by author(s) and year of publication as in the Publication Manual of the American Psychological Association.